Lecture Notes in Electrical Engineering 244

Jiadong Sun · Wenhai Jiao
Haitao Wu · Chuang Shi
Editors

China Satellite Navigation Conference (CSNC) 2013 Proceedings

Satellite Navigation Signal System, Compatibility & Interoperability • Augmentation & Integrity Monitoring • Models & Methods

Springer

Lecture Notes in Electrical Engineering

Volume 244

For further volumes:
http://www.springer.com/series/7818

Jiadong Sun · Wenhai Jiao
Haitao Wu · Chuang Shi
Editors

China Satellite Navigation Conference (CSNC) 2013 Proceedings

Satellite Navigation Signal System, Compatibility & Interoperability • Augmentation & Integrity Monitoring • Models & Methods

Springer

Editors
Jiadong Sun
China Aerospace Science and Technology Corporation
Chinese Academy of Sciences
Beijing
People's Republic of China

Wenhai Jiao
China Satellite Navigation Office
Beijing
People's Republic of China

Haitao Wu
Navigation Headquarters
Chinese Academy of Sciences
Beijing
People's Republic of China

Chuang Shi
Wuhan University
Wuhan
People's Republic of China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
ISBN 978-3-642-37403-6 ISBN 978-3-642-37404-3 (eBook)
DOI 10.1007/978-3-642-37404-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013938171

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

China's BeiDou Navigation Satellite System (BDS) has been independently developed, which is similar in principle to global positioning system (GPS) and compatible with other global satellite navigation systems (GNSS). The BeiDou will provide highly reliable and precise positioning, navigation, and timing (PNT) services as well as short-message communication for all users under all-weather, all-time, and worldwide conditions.

Since BeiDou Navigation Satellite System provided the test run services on December 27, 2011, more than 6 satellites have been successfully launched in 4 times with large improvements in system coverage, constellation robustness and positioning accuracy. Currently, all in-orbit satellites and ground systems run well, which meet the design requirements through the testing and evaluation of various user terminals. After the news conference announced the Full Operational Capability (FOC) of BeiDou Navigation Satellite System for China and surrounding area on December 27, 2012, the BeiDou Navigation Satellite System officially starts to provide continuous passive positioning, navigation and timing services as well as active positioning, two-way timing and short message communication services.

China Satellite Navigation Conference (CSNC) is an open platform for academic exchanges in the field of satellite navigation. It aims to encourage technological innovation, accelerate GNSS engineering and boost the development of the satellite navigation industry in China and in the world.

The 4th China Satellite Navigation Conference (CSNC 2013) is held on May 13–17, 2013, Wuhan, China. The theme of CSNC 2013 is BeiDou Application—Opportunities and Challenges, which covers a wide range of activities, including technical seminars, academic exchange, forum, exhibition, lectures as well as ION panel. The main topics are as:

1. BeiDou/GNSS Navigation Applications
2. Satellite Navigation Signal System, Compatibility and Interoperability
3. Precise Orbit Determination and Positioning
4. Atomic Clock Technique and Time–Frequency System

5. Satellite Navigation Augmentation and Integrity Monitoring
6. BeiDou/GNSS Test and Assessment Technology
7. BeiDou/GNSS User Terminal Technology
8. Satellite Navigation Models and Methods
9. Integrated Navigation and New Methods

The proceedings have 181 papers in nine topics of the conference, which were selected through a strict peer-review process from 627 papers presented at CSNC 2013.

We thank the contribution of each author and extend our gratitude to over 100 referees and 36 session chairmen who are listed as members of editorial board. The assistance of CSNC 2013's organizing committees and the Springer editorial office is highly appreciated.

Jiadong Sun
Chair of CSNC 2013

Editorial Board

The 4th China Satellite Navigation Conference (CSNC 2013)

Scientific Committee

Chairman:
Jiadong Sun, China Aerospace Science and Technology Corporation

Vice-Chairman:
Rongjun Shen, China
Jisheng Li, China
Qisheng Sui, China
Zuhong Li, China Academy of Space Technology
Shusen Tan, Beijing Satellite Navigation Center, China

Executive Chairman:
Jingnan Liu, Wuhan University
Yuanxi Yang, China National Administration of GNSS and Applications
Shiwei Fan, China

Committee Members (By Surnames Stroke Order):
Xiancheng Ding, China Electronics Technology Group Corporation
Qingjun Bu, China
Liheng Wang, China Aerospace Science and Technology Corporation
Yuzhu Wang, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences
Guoxiang Ai, National Astronomical Observatories, Chinese Academy of Sciences
Shuhua Ye, Shanghai Astronomical Observatories, Chinese Academy of Sciences
Zhaowen Zhuang, National University of Defense Technology
Qifeng Xu, PLA Information Engineering University
Houze Xu, Institute of Geodesy and Geophysics, Chinese Academy of Sciences
Guirong Min, China Academy of Space Technology
Xixiang Zhang, China Electronics Technology Group Corporation
Lvqian Zhang, China Aerospace Science and Technology Corporation

Junyong Chen, National Administration of Surveying, Mapping and Geoinformation
Benyao Fan, China Academy of Space Technology
Dongjin Luo, China
Guohong Xia, China Aerospace Science and Industry Corporation
Chong Cao, China Research Institute of Radio Wave Propagation (CETC 22)
Faren Qi, China Academy of Space Technology
Sili Liang, China Aerospace Science and Technology Corporation
Shancheng Tu, China Academy of Space Technology
Rongsheng Su, China
Zhipeng Tong, China Electronics Technology Group Corporation
Ziqing Wei, Xi′an Institute of Surveying and Mapping

Organizing Committee

Secretary General:
Haitao Wu, Navigation Headquarters, Chinese Academy of Sciences

Vice-Secretary General:
Wenhai Jiao, China Satellite Navigation Office
Tao Peng, Academy of Satellite Application, China Aerospace Science and Technology Corporation
Wenjun Zhao, Beijing Satellite Navigation Center
Chuang Shi, Wuhan University
Weina Hao, Navigation Headquarters, Chinese Academy of Sciences

Committee Members (By Surnames Stroke Order):
Qun Ding, Xi'an Institute of Navigation Technology
Miao Yu, Beijing Institute of Space Science and Technology Information
Yinhu Ma, Navigation Satellite Systems Engineering Center, China Academy of Aerospace Electronics Technology
Xiuwan Chen, Institute of Digital China, Peking University
Mingquan Lu, Tsinghua University
Guangzhou Ouyang, Academy of Opto-Electronics, Chinese Academy of Sciences
Shuanggen Jin, Shanghai Astronomical Observatory, Chinese Academy of Sciences
Xiang'an Zhao, China Defense Science and Technology Information Center
Hong Yuan, Navigation Headquarters, Chinese Academy of Sciences
Yamin Dang, Chinese Academy of Surveying and Mapping
Zhong Dou, National Time Service Center, Chinese Academy of Sciences

Contents

Part I
Satellite Navigation Signal System, Compatibility and Interoperability

Chapter 1
A Research of GNSS RF Interference Assessment Method

Yao Wang and Baoguo Yu

Abstract The Radio Frequency Interference (RFI) was one of the main problems faced by Global Navigation Satellite System (GNSS), However, there were several limits in current RFI assessment method. Based on Multilevel Coded Spreading Symbols (MCS) Signal, a generic signal baseband model was constructed. On the basis of research of A/D model, interference coefficient model, effective C/N_0 model, receiving sensitivity model and ranging error model including interference, a improved RFI evaluation method was proposed. Meanwhile, some important models in evaluation method were analyzed by simulation, and some useful results were got. The achievement by this paper would guide the frequency coordination negotiation between Compass and other GNSSs.

Keywords Radio frequency interference · GNSS · Compatibility · MCS

1.1 Introduction

With the construction of Galileo system and global application of Chinese BD system, by 2020, the world will form the global navigation satellite systems of GPS, GLONASS, Galileo, BD and some regional or enhanced system. Then, there is more than 120 navigation satellites in the space, and more than 30 kinds of signals will be launched. While the satellite navigation frequency band allocated by ITU is limited, this inevitably brings the interference between systems.

Y. Wang (✉)
The 54th Institute of China Electronic Technology Corporation, Shijiazhuang, Hebei 050081, China
e-mail: wy_hrbeu@163.comwy_hrbeu@sohu.com

Y. Wang · B. Yu
Hebei Satellite Navigation Technology and Equipment Engineering Technology Research Centre, Shijiazhuang, Hebei 050081, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_1,

Meanwhile, there are also lots of unintentional interference, which will degrade the system performance. A coordination methodology for RNSS inter-system interference assessment was proposed by International Telecommunication Union (ITU), but there are some limits in it [1].

In recent years, many researchers at home and abroad have studied RFI assessment [2–4], but the problem is not considered comprehensive, mainly including as follows: (1) not considering modern signal with pilot and data channel; (2) not considering the effect of sampling and quantization of receiver front; (3) lacking of solving method of interference threshold; (4) not considering the effect of RFI onto range accuracy. In this paper, based on MCS signal, a generic signal baseband model is constructed. And a improved RFI assessment methodology is proposed. Finally, some important models in evaluation method are analyzed by simulation, and some useful results are got.

1.2 Signal Model

The interference analysis of GNSS is related to signal power spectral density. A generic signal baseband spectrum domain model is constructed based on MCS signal as baseline. MCS could be expressed as MCS $([m_k], n)$. Where, $[m_k]$ represents multilevel coded sequence and n represents PRN rate, i.e., $n \times 1.023$ MHz. MCS may derive a number of modulation modes, e.g., Binary Coded Symbol (BCS), Composite BCS (CBCS), Sinusoidal Multilevel Coded Symbol (SMCS), etc. The relation among them is as shown in the figure below (Fig. 1.1).

Assuming that T_c is PRN code chip length, $T_s = T_c/K$ is one multilevel coded width, the chip waveform of MCS signal could be expressed as follows:

$$\phi(t) = \sum_{k=1}^{K} m_k p\left(\frac{t - kT_s}{T_s}\right) \tag{1.1}$$

where

$$p\left(\frac{t}{T_s}\right) = \begin{cases} 1 & 0 \le t \le T_s \\ 0 & else \end{cases} \tag{1.2}$$

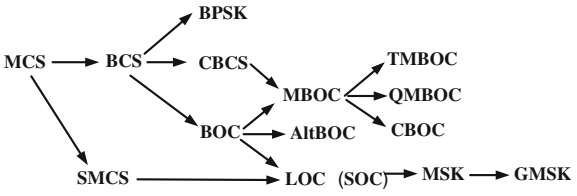


Fig. 1.1 Derivation relationship of various chip waveforms

Then, the Fourier transformation of chip waveform may be expressed as follows:

$$\Phi(f) = F\left\{\sum_{k=1}^{K} m_k p\left(\frac{t - kT_s}{T_s}\right)\right\} = \frac{\sin(\pi f T_s)}{\pi f} e^{j\pi f T_s} \sum_{k=1}^{K} m_k e^{-j2\pi f k T_s} \tag{1.3}$$

The power spectrum of long code signal could be expressed as follows

$$G_{s,l}(f) = \frac{1}{T_c}\Phi^2(f) \tag{1.4}$$

Meanwhile, the power spectrum of short code signal can be expressed as follows [5]

$$\begin{aligned} G_{s,s}(f) &= \frac{T_d}{(NT_c)^2}\Phi^2(f)X_{code}^2(f)\sum_{k=-\infty}^{\infty} \sin c^2\left(\pi T_d\left(f - \frac{k}{NT_c}\right)\right) \\ &= \frac{T_d}{N^2 T_c} G_{s,l}(f)X_{code}^2(f)\sum_{k=-\infty}^{\infty} \sin c^2\left(\pi T_d\left(f - \frac{k}{NT_c}\right)\right) \end{aligned} \tag{1.5}$$

While, T_d is single message symbol length, N is PRN code period, $X_{code}(f)$ is code transformation. Let expression (1.3) be used in expression (1.4), then

$$\begin{aligned} {}_{MCS,l}(f) &= \frac{\sin^2(\pi f T_s)}{T_c(\pi f)^2}\left|\sum_{k=1}^{K} m_k e^{-j2\pi f k T_s}\right|^2 \\ &= T_s \sin c^2(\pi f T_s)\left\{\frac{1}{K}\left|\sum_{k=1}^{K} m_k e^{-j2\pi f k T_s}\right|^2\right\} \\ &= G_{subchip}(f) \cdot G_{\mathrm{mod}}(f) \end{aligned} \tag{1.6}$$

While, $G_{subchip}(f)$ represents the unit waveform of every code symbol and $G_{\mathrm{mod}}(f)$ represents the modulation factor. The PSD of all signals derived from MCS could be expressed product of the two parts, as shown in Table 1.1.

1.3 RFI Assessment Method

1.3.1 ITU Assessment Method

For the assessment methods used by the ITU, refer to the reference documents [2], known as "the 1831 Proposal" for short. In the proposal, equivalent C/N_0 is used as the assessment factor of GNSS RFI, as expressed below [1, 6]

Table 1.1 The PSD expression of several GNSS signals

Modulation	PSD expression	
	$G_{\text{subchip}}(f)$	$G_{\text{mod}}(f)$
BPSK	$T_s \sin c^2(\pi f T_s)$	$\frac{1}{K}\frac{\sin^2(\pi f T_c)}{\sin^2(\pi f T_s)}$
BOC_s (m, n)	$T_s \sin c^2(\pi f T_s)$	$1/K \begin{cases} \sin^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is even} \\ \cos^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is odd} \end{cases}$
BOC_c (m, n)	$\frac{4\sin^4(\pi/2fT_s)}{\pi^2 f^2 T_s}$	$1/K \begin{cases} \sin^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is even} \\ \cos^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is odd} \end{cases}$
SOC (m, n)	$\frac{1}{T_s}\frac{8\pi^2\cos^2(\pi f T_s)}{\left(\pi^2 - 4\pi^2 f^2 T_s^2\right)^2}$	$1/K \begin{cases} \sin^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is even} \\ \cos^2\left(\frac{\pi f}{f_c}\right) \Big/ \cos^2\left(\frac{\pi f}{K f_c}\right), K \text{ is odd} \end{cases}$

$$\left(\frac{C}{N_0}\right)_{eff} = \frac{C}{\eta N_0 + I_{\text{inter}} + I_{\text{intra}} + I_{\text{ext}}} \tag{1.7}$$

Where, C represents the signal power processed by correlator; η represents the active power factor characterizing the percentage of desired signal power in the total power; I_{inter}, I_{intra} and I_{ext} represent the spectral density of equivalent thermal noise power brought in by intrasystem signal, intersystem signal and external signal respectively (based on the output of instant channel correlator). The change of equivalent C/N_0 is used to assess the interference level in case of any. In the description of the proposal, the following points are worth for further discussion: (1) the model of equivalent C/N_0 is derived on the basis of the SINR of the digital signal instant channel output, therefore, the effect of sampling, quantization and correlation loss should be considered in the C/N_0 model of received desired signal; (2) modernized signals are all based on the pilot frequency and data dual channel, and no such C/N_0 model is available so far; (3) the effect of RFI to lead/lag channel should be taken into consideration; (4) interference might cause acquisition loss, unlocked loop or message demodulation failure, the sensitivity of receiver in every processing course should be considered; (5) under the condition of normal positioning, the RFI result will directly affect the ranging accuracy, and no such type of model is available; finally, the effect of PRN and data to assessment method should also be taken into account (in association with the lumping gain factor).

1.3.2 Method Improvement

1.3.2.1 Model of Receiving and Processing Loss

The effect of receiver front end to C/N_0 is mainly the sampling/quantizing process. Higher sampling rate and more number of quantization bits have less effect to C/N_0. Quantization would generate noise, which would lower the SNR. Oversampling in the sampling process would improve the SNR. Jitter might also cause noise, which however gives very little effect to the SNR. The comprehensive effect of receiver processing to C/N_0 are expressed as follows:

$$L_{loss} = L_{sampling} + L_{digital} + L_{corr} \tag{1.8}$$

where,

$$L_{sampling} = 10 \times \lg\left(\frac{f_{sampling}}{2B}\right) \tag{1.9}$$

$$L_{digital}(D) = 10\lg\left(\frac{2}{\pi}\frac{\left[1 + 2\sum_{i=1}^{2^{D-1}-1}\exp\left\{-\frac{(i/A_g)^2}{2\sigma_{IF}^2}\right\}\right]^2}{1 + 8\sum_{i=1}^{2^{D-1}-1} i \cdot erfc\left(\frac{i/A_g}{\sqrt{2}\sigma_{IF}}\right)}\right) \tag{1.10}$$

$$L_{corr} = P_{\substack{CCF \\ ref}} - P_{\substack{CCF \\ real}} \tag{1.11}$$

In Expression (1.9), $L_{sampling}$ represents the C/N_0 loss caused by sampling; $f_{sampling}$ represents the sampling frequency of AD; B represents the signal bandwidth. In Expression (1.10), $L_{digital}$ represents the C/N_0 loss caused by quantization, D represents the quantizing bit; A_g is the signal gain through AGC step; and σ_{IF} is the noise RMS.

L_{corr} represents the correlation loss [7] resulting from imperfect signal. There are two main causes to correlation loss: (a) a number of signal components are multiplexed at the same carrier frequency; (b) band limited noise and signal distortion at the same carrier frequency.

1.3.2.2 Interference Coefficient Model

In the 1831 Proposal, the spectral density of equivalent thermal noise power is computed on the basis of spectrum splitting coefficient [SSC, refer to Expression (1.12)]. However, by analyzing the code tracking process, it is found that the spectral density of equivalent thermal noise power depends on the code tracking spectrum sensitivity coefficient [8] [CT_SSC, refer to Expression (1.13), where Δ

represents the chip delay between lead branch and lag branch]. If Δ is very small, μ_{ls} may be known as code tracking spectrum splitting coefficient.

$$\kappa_{ls} = \frac{\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} G_l(f+\Delta f)G_s(f)df}{\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} G_s(f)df} \tag{1.12}$$

$$\mu_{ls} = \frac{\int_{-\beta_r/2}^{\beta_r/2} G_l(f+\Delta f)G_s(f)\sin^2(\pi f\Delta)df}{\int_{-\beta_r/2}^{\beta_r/2} G_s(f)\sin^2(\pi f\Delta)df} \tag{1.13}$$

As shown in the two expressions above, the difference between the two expressions is one $\sin^2(\pi f\Delta)$; therefore, be sure to take the maximum when assessing the system compatibility, which may be known as interference coefficient.

$$\xi_{ls} = \max\{\mu_{ls}, \kappa_{ls}\} \tag{1.14}$$

1.3.2.3 I/Q Equivalent C/N_0 Model

The modernized GNSS signals are all compound signals, which consist of data branch and pilot branch. If only one branch is acquired, signal power loss will be incurred. This is very unfavorable for the acquisition process that is sensitive to signal power; therefore, a good method must combine both data branch and pilot frequency branch. One of the most classical methods is combined non-coherent acquisition method [7, 9]. In this method, the correlator output SINR is as follows:

$$\rho_{cnc} \approx \frac{\left(C_D R_{D,\phi}^2 + C_P R_{P,\phi}^2\right)\sin c^2(\pi\Delta f_D T_{co})}{\frac{1}{T_{co}}\int_{T_{co}} R_w(\tau)\left[R_{D,\phi}(\tau)+R_{P,\phi}(\tau)\right]e^{j2\pi\Delta f\tau}d\tau} + 1 \tag{1.15}$$

The equivalent C/N_0 is derived as follows:

$$\left(\frac{C}{N_0}\right)_{eff} = \frac{C\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}}\left(G_{P,s}(f)+G_{D,s}(f)\right)df}{N_0\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}}\left(G_{P,s}(f)+G_{D,s}(f)\right)df + C_l\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} G_l(f+\Delta f)\left(G_{P,s}(f)+G_{D,s}(f)\right)df} \tag{1.16}$$

For the definitions of the correlation variables in the expression above, refer to the reference documents [10]. If the interference also consists of two branches I/Q, then

$$SSC = SSC(s_D, l_D) + SSC(s_D, l_P) + SSC(s_P, l_D) + SSC(s_P, l_P) \tag{1.17}$$

where, s_D and s_P represent the data channel and the pilot channel of desired signal respectively; l_D and l_P represent the data channel and the pilot channel of interference signal respectively.

1.3.2.4 Receiver Sensitivity

The receiver sensitivity may assure the minimum C/N_0 of signal receiver and satisfy the requirements of signal acquisition, tracking and data demodulation [11]. The assessment factor for the acquisition process is acquisition time. The factors affecting initial acquisition time include detection probability, false alarm probability, physical structure of receiver, etc. The assessment factor for code tracking process is code-tracking error. The factors affecting code-tracking error include thermal noise, dynamic stress, tracking loop construction, etc. The assessment factor for carrier tracking is mean cycle slip time and carrier tracking error, which are mainly related to carrier loop design, oscillator stability and external environment. The assessment factor for data demodulation is Frame Error Rate (FER). Work out the minimum C/N_0 values for the four receiver processing steps separately and take the maximum value, which is the minimum C/N_0 that satisfies the signal receiving requirement, as shown in the figure below (Fig. 1.2).

1.3.2.5 Effect of RFC to Ranging

With reference documents [10] and the tracking loop model, the code tracking error of coherent DLL is as follows:

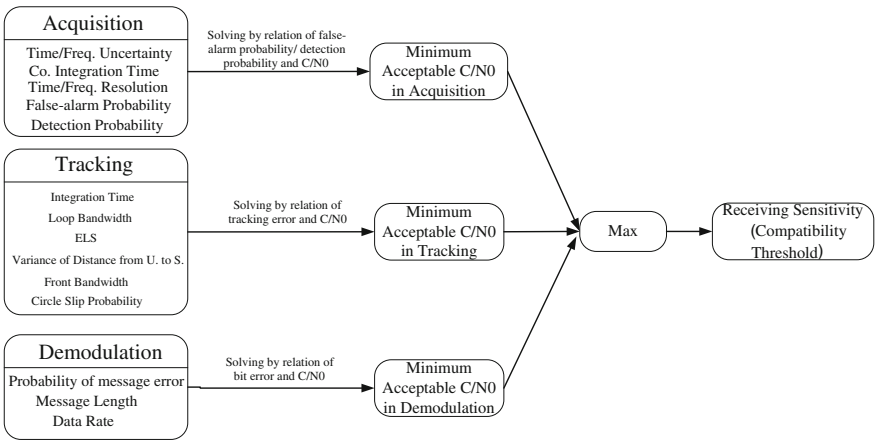


Fig. 1.2 Schematic diagram of receiver sensitivity estimation

$$\sigma_{CELP}^2 = \sqrt{B_n \frac{\int_{-\beta_r/2}^{\beta_r/2} \left[(C/N_0)^{-1} + (C/C_l)^{-1} G_l(f+\Delta f)\right] G_s(f) \sin^2(\pi f \delta T) df}{4\pi^2 \left(\int_{-\beta_r/2}^{\beta_r/2} f G_s(f) \sin(\pi f \delta T) df\right)^2}} \tag{1.18}$$

When receiving pilot/data separated signal, three carrier tracking modes are considered: data-channel tracking demodulation without pilot frequency aid, data channel demodulation aided with pilot-channel tracking and data/pilot channel combined tracking demodulation. The thermal noise carrier tracking errors are respectively as follows in the above three modes:

$$\sigma_D = \frac{360}{2\pi} \sqrt{\frac{B_L \int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} G_w(f+\Delta f) G_s(f) df}{a_D C \left(\sin c(\pi \Delta f_D T_{co}) \int_{-\beta_r/2}^{\beta_r/2} G_s(f) e^{j2\pi f \Delta\tau} df\right)^2} \hbar \left(1 + \frac{\int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} G_w(f+\Delta f) G_s(f) df}{2T_{co} a_D C \left(\sin c(\pi \Delta f_D T_{co}) \int_{-\beta_r/2}^{\beta_r/2} G_s(f) e^{j2\pi f \Delta\tau} df\right)^2}\right)} \tag{1.19}$$

$$\sigma_P \approx \frac{360}{2\pi} \sqrt{\frac{B_L \int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} N_0 G_s(f) df + \int_{-\frac{\beta_r}{2}}^{\frac{\beta_r}{2}} C_l G_l(f+\Delta f) G_s(f) df}{a_P C \left(\int_{-\beta_r/2}^{\beta_r/2} G_s(f) df\right)^2}} \text{ (Deg.)} \tag{1.20}$$

$$\sigma_C = \sqrt{\frac{\sigma_D^2 \cdot \sigma_P^2}{\sigma_D^2 + \sigma_P^2}} \text{ (Deg.)} \tag{1.21}$$

1.3.3 Summary of Assessment Method

A improved RFI assessment method is proposed based on the analysis above, as shown in Fig. 1.3. The key of assessing satellite navigation RFI is to solve three parameters, as follows: (1) the desired signal power before A/D; (2) the interference coefficient; (3) the receiver sensitivity. Among them, the attenuation amount of transmission link should be considered in desired signal power; the Doppler, PRN code, message and filter should be considered in interference coefficient. The parameters of the improved method finally got are degression of effective C/N_0, ranging error, whether the desired signal can be received normally. Finally, the three parameters above should be comprehensive evaluated to judge whether it meets the system requirements.

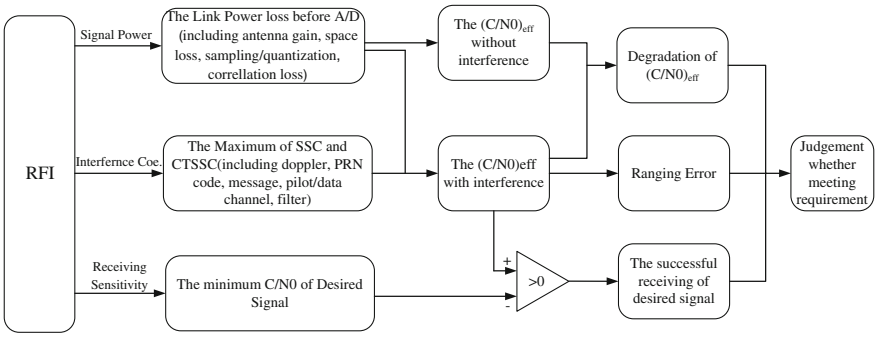


Fig. 1.3 The improved RFI assessment method

1.4 Simulation and Validation

The RFI method employed by ITU is improved by last chapter. And some important models in it are analyzed by simulation in this chapter as follows.

1.4.1 The Effect of Quantification Onto C/N$_0$

The quantification of the receiver would cause the loss of energy of signal power, and it would further cause C/N$_0$ of desired signal. In order to provide the point of view, GPS C/A signal is illustrated in the following figure. The navigation signal is buried in the noise, so, the AGC of receiver is controlled by the noise power. The figure shows that, the signal-to-noise ratio isn't increased too much as $A_g\sigma_{IF}$ increase at the appropriate point (Fig. 1.4) (Table 1.2).

1.4.2 Analysis of Interference coefficient

There is different between SSC and CTSSC for different signals. And they are simulated as follow while the desired and interference signals are BPSK and BOC modulation respectively. And Tables 1.3 and 1.4 are got. Accordingly, the cures are shown in appendix.

As shown in Table 1. 3, when the interference signal is BPSK modulation, the SSC is larger than CTSSC only in the situation that BPSK or BOC (1,1) is the interference signal.

As can be seen in Table 1.4, when the desired signal is BOC, CTSSC is larger than SSC, In spite of the interference signal is BPSK or BOC.

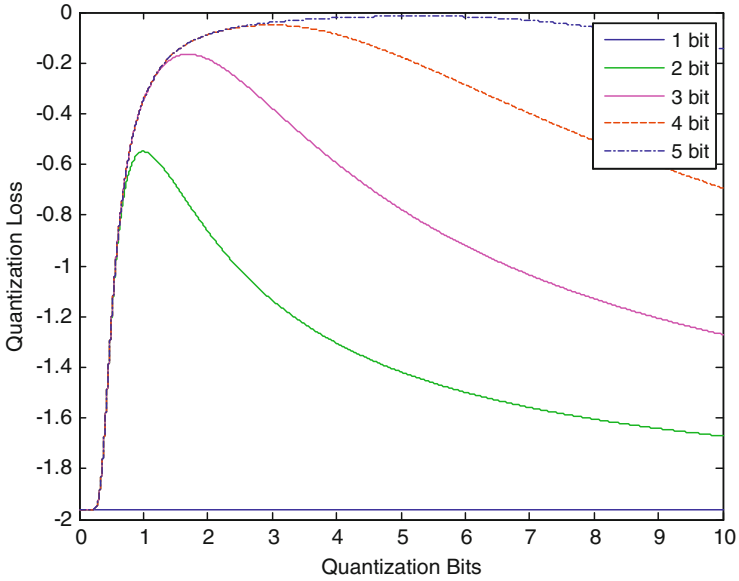


Fig. 1.4 The relationship between quantization bits and quantization loss

Table 1.2 The minimum loss of different quantization bits

Quantiztion bits (bit)	The optimal AGC gain (A_g)	The minimum loss (dB)
1	———	−1.9612
2	$1.005/\sigma_{IF}$	−0.5495
3	$1.707/\sigma_{IF}$	−0.1657
4	$2.984/\sigma_{IF}$	−0.0504
5	$5.315/\sigma_{IF}$	−0.0152

Table 1.3 The difference between SSC and CTSSC (1)

Desired signal	Interference signal (CTSSC–SSC)/dB			
	BPSK	BOC (m, 1)	BOC (m, 2)	BOC (m, 3)
BPSK(1)	<0	>0	>0	>0
BPSK(2)	<0	>0	>0	>0
BPSK(3)	<0	>0 (<0, m = 1)	>0	>0
BPSK(4)	<0	>0 (<0, m = 1)	>0	>0
BPSK(10)	<0	>0 (<0, m = 1)	>0	>0

Table 1.4 The difference between SSC and CTSSC (2)

Desired signal	Interference Signal (CTSSC–SSC)/dB			
	BPSK	BOC (m, 1)	BOC (m, 2)	BOC (m, 3)
BOC (1,1)	>0	>0	>0	>0
BOC (2,1)	>0	>0	>0	>0
BOC (3,1)	>0	>0	>0	>0
BOC (4,1)	>0	>0	>0	>0
BOC (10,1)	>0	>0	>0	>0

1.4.3 The Effect of RFI Onto Ranging

Firstly, the influence of RF interference on the ranging accuracy of the code tracking is analyzed. If the receiver front-end bandwidth is 24 MHz, bandwidth of the code loop is 2 Hz, carrier to noise ratio of desired signal is 44 dB-Hz, the modulation is BPSK (1) and BOC (1,1), Fig. 1.5 shows the desired signal is a BPSK (1) and Fig. 1.6 shows the desired signal is a BOC. It can be seen from the figure that ranging error caused by interference don't change obviously with ELS of the loop the when ISR is small. However, when ISR is large, ranging error caused by the interference change with the interval of loop significantly. It should be noted that, whether the ISR is small or large, and whatever the interference signal and desired signal is modulated, ranging error has the same trend with ELS (Fig. 1.7).

For carrier tracking, considering there is no data, namely, the formula (1.20), the loop bandwidth is 10 Hz, desired signal is BPSK (1), the interference signal is BPSK (1) and BOC (1,1).The tracking error caused by ISR and Doppler shift are shown in Fig. 1.8. It can be seen that the tracking error increases as the ISR and Doppler become large, and changes little with tracking error. The analysis of the interference between GPS L1C/A and Galileo E1OS [modulated by BOC (1,1)] are shown in Fig. 1.9, it can been seen that results make a difference with the ideal

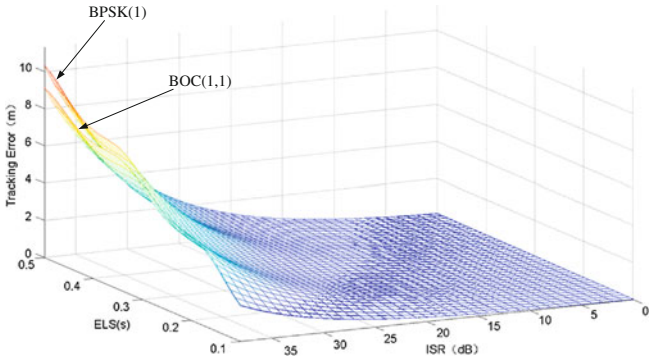


Fig. 1.5 The code tracking error caused by RFI [desired signal is BPSK (1)]

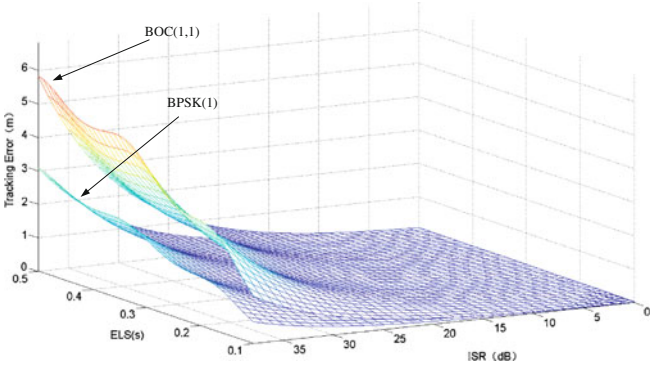


Fig. 1.6 The code tracking error caused by RFI [desired signal is BOC (1,1)]

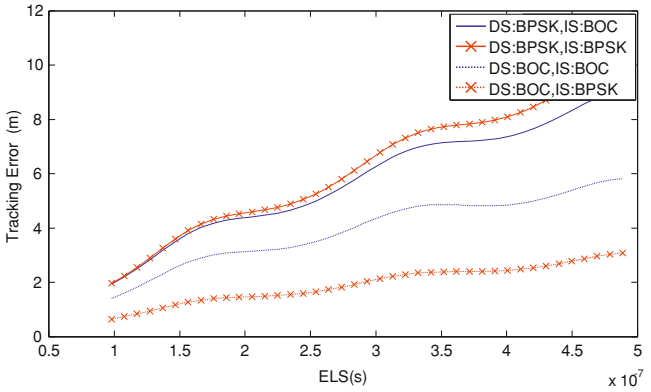


Fig. 1.7 The code tracking error caused by RFI

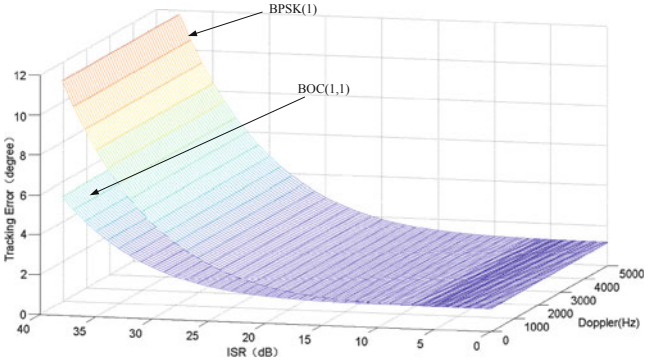


Fig. 1.8 The tracking error caused by RFI [desired signal is BPSK (1)]

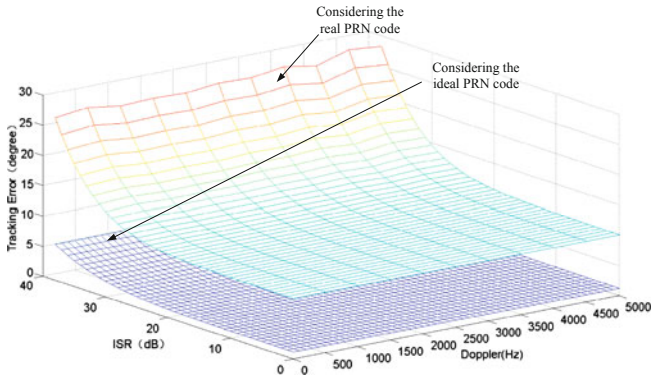


Fig. 1.9 The tracking error caused by RFI (desired signal is GPS C/A)

case, so it should consider the impact of the pseudo-code when analyzing the mutual interference of the civilian signal

It should be pointed out that the interferences between BPSK (1) and BOC (1,1) are analyzed, although, the conclusions are universal.

1.5 Conclusion

The RFI could cause performance degradation of the navigation system. And, the RFI includes not only the intersystem interference of GNSS, but also other unintentional interference within navigation bands. In order to solve the RFI within GNSS band, ITU released a RFI assessment method, also known as the 1831 recommendation. This paper has made two improvements: on the one hand, on the other hand the universal expression of power spectral density of signal in theory is derived, inclining long code and short code; on the other hand, a improved RFI assessment method is proposed based on 1831 recommendation. Finally, simulation has been made about the effect of quantization onto C/N_0, the interference coefficient, and the effect of RFI onto ranging. The conclusions are got as follows:

1. The influence of quantization onto GNSS compatibility can not be ignored.
2. The SSC is larger than CTSSC in the situation of desired signal and interference signal are both BPSK modulation ones.
3. The code tracking error caused by different spreading spectrum signals differs from ELS, the tendency is the same;
4. The Doppler frequency difference between the interference signal and the desired signal could cause the fluctuation of the carrier tracking error in the conditions of the short code;
5. Ranging error caused by the interference increases exponentially with the ISR.

References

1. ITU (2007) A coordination methodology for RNSS inter-system interference assessment[EB/OL]. http://webs.uvigo.es/servicios/biblioteca/uit/rec/M/R-REC-M.1831-0-200710-I!!PDF-E.pdf
2. Zhang J, Ma J, Chen G (2012) Research on compatibility analysis methods for global satellite navigation systems. J Spacecraft TT and C Technol 2(31):73–76
3. QiBing X, Lixin Z, Yansong M (2011) Analysis of the compatibility of the GNSS signal system. Space Electron Technol 2(1–4):42
4. Ran Y (2011) Research on modulation and frequency compatibility of GNSS signals. Huazhong University of Science and Technology, Wuhan
5. Wang Y, Yu BG, Luo XZ (2010) More accurate methodology for assessing multi-mode GNSS compatibility. Syst Eng Electron 32(6):1305–1308
6. Rodriguez JAA (2008) On generalized signal waveforms for satellite navigation. University FAF Munich, Munich
7. Rapisarda M, Hannes D, Burger T et al (2009) The off-line SIS quqlity analysis tool: design, development and validation. ION GNSS 2009[C]. Alexandria, VA: ION, 2899–2910
8. Wang Y, Yang S, Yu B (2010) Research on compatibility analysis between GPS and galileo systems based on code tracking. Acta Armamentarii 31(12):1680–1685
9. Dierendonck AJV, Spilker JJ (1999) Proposed civil GPS signal at 1176.45 MHz. In: phase/Quadrature Codes at 10.23 MHz Chip Rate. ION Annual Meeting. Cambridge, MA, 761–770
10. Wang Y, Yu B, Luo X (2011) Theoretical research and simulation on GNSS compatibility. The second China Satellite Navigation Conference 2011
11. Bastide F, Julien O, Macabiau C et al (2002) Analysis of L5/E5 acquisition, tracking and data demodulation thresholds. ION GPS/GNSS, Portland, pp 2196–2207

Chapter 2
A Quasi-Cyclic LDPC Code for GNSS Signal

Yi Yang, Changjian Liu and Xiaoqing Zhang

Abstract Considering that modernized GPS L1C signal utilizes random constructed LDPC as the FEC coding scheme, it's a waste of storage resources because of the irregularity of the parity-check matrix, though owning superb error-control performance. This paper proposed a constructing algorithm for Quasi-Cyclic LDPC code which is suitable for GNSS signal, it avoids the existence of short cycle with the length of 4 in the Tanner graph from the beginning, by limiting the cross-correlation values of sparse sequences. Better performance can be achieved by making the parity-check matrix sparser. Using 802.11n protocol as a reference, parity-check part is double diagonal, which makes the encoding process possible directly with parity-check matrix. Simulation results show this LDPS code performs slightly better than what GPS L1C uses, lower encoding and decoding complexity makes it more suitable for GNSS.

Keywords Satellite navigation · Low-density parity check (LDPC) code · Parity-check matrix · Quasi-cyclic code

2.1 Introduction

Growing number of new ideas are prominently emerging with the development of GNSS signal design [1], including the introducing of forward error correction (FEC) coding such as convolutional coding and LDPC code for the purpose of improving the signal's robustness. Conventional GPS C/A signal utilizes Hamming code to make itself possible to detect errors, but not able to forward error control. Modernized GPS signal generally uses convolutional code as the

Y. Yang (✉) · C. Liu · X. Zhang
Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: yangyi@hust.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_2,

FEC code with the coding efficiency of 1/2, while GPS L1C uses LDPC code which results in higher coding gain in contrast with convolutional code as an exception.

Subframe 2 and subframe 3 of GPS L1C are separately encoded using rate 1/2 LDPC codes, as a result, there are 1,200 coded bits for subframe 2 and 548 coded bits for subframe 3 [2]. L1C constructs the parity-check matrices randomly which makes them superb at error correcting when codeword is long, but on the other hand, encoding and hardware realization complexities are substantially increased due to the long codeword and irregularity of parity-check and generator matrices. Although GPS uses lower-triangular-matrix-based LDPC code, RU algorithm [3] gives out an encoding scheme with approximately linear complexity, randomly constructed parity-check matrices and relatively high average row weights still make the requirements of storage and calculation a demanding problem.

Structured LDPC code, especially Quasi-cyclic LDPC code is proposed for above reason, which owns lower encoding complexity and makes it easier for hardware realization. Shanbao et al. [4] proposes an algorithm to obtain quasi-cyclic parity-check matrix by finding a series of so-called sparse sequences, but the method for removing 6-cycle in the Tanner graph can just apply to the situation with small number of sequences, what's more, none of the sub-matrices is zero matrix and that has an negative effect on error correction performance. Combining with relevant theory, [5] discusses how to remove 4-cycles and 6-cycles, but not able to explain why there appear zero sub-matrices on the left side of parity-check matrix.

Based on the work of [4], this paper generates several sparse sequences make sure the cross-correlation values between any two arbitrary generated ones are no more than 1, by this way a parity-check matrix is obtained which is dual-diagonal by borrowing ideas from protocol IEEE 802.11n. Finally, better performance can be achieved by setting some of the sub-matrices zero randomly.

2.2 Quasi-Cyclic LDPC Code and its Encoding Algorithm

2.2.1 Characteristics of the Parity-Check Matrix

The characteristic of QC-LDPC code is that there is a number z, we can divide any available codeword with the length of $n * z$ into n pieces of sub-codewords and the resulting codeword is still available if we cyclic shift each sub-codeword by a common length separately, so the corresponding parity-check matrix consists of several sub-matrices with the same size $(z \times z)$, assume each sub-matrix is a circulant permutation matrix or a zero matrix, then the parity-check matrix can be expressed as **B**:

$$\mathbf{B} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{bmatrix} \tag{2.1}$$

which we call *basis matrix*. For any $p_{i,j}\left(-1 \le p_{i,j} < z\right)$, it represents a sub-matrix derived by cyclic shifting identity matrix $\mathbf{I}(0)$ by $p_{i,j}$ steps if $p_{i,j} \ge 0$, and represents zero matrix if $p_{i,j} = -1$. It's clearly seen that the performance of codeword is totally determined by basis matrix once codeword length and size of sub-matrix z are given, so it should be meticulously designed and optimized. Wang et al. [6] gives out a theorem on judging if short cycle exists in the parity-check matrix by probing its basis matrix.

Since parity-check matrix can be expressed by its basis matrix which has lower dimension, storage requirements are significantly decreased. From the standpoint of hardware, QC-LDPC encoding process can be easily accomplished by circulating registers and trade off time complexity against space complexity [7].

Using dual-diagonal structure, the parity-check matrix can be directly used to encode thus there is no need to calculate generator matrix. Considering the case coding efficiency is 1/2, which means $m = n/2$, basis matrix has the following form as (2.2) shows

$$\begin{aligned} \mathbf{B} &= [\mathbf{B_S}|\mathbf{B_P}]_{m\times 2m} \\ &= \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,m} & 1 & 0 & -1 & -1 & \cdots & -1 & -1 \\ p_{2,1} & p_{2,2} & \cdots & p_{2,m} & 0 & 0 & 0 & -1 & \cdots & -1 & -1 \\ p_{3,1} & p_{3,2} & \cdots & p_{3,m} & -1 & -1 & 0 & 0 & \cdots & -1 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{i,1} & p_{i,2} & \cdots & p_{i,m} & -1 & -1 & -1 & -1 & \cdots & -1 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{m-1,1} & p_{m-1,2} & \cdots & p_{m-1,m} & -1 & -1 & -1 & -1 & \cdots & 0 & 0 \\ p_{m,1} & p_{m,2} & \cdots & p_{m,m} & 1 & -1 & -1 & -1 & \cdots & -1 & 0 \end{bmatrix} \end{aligned} \tag{2.2}$$

$\mathbf{B_S}$ corresponds to parity-check coefficients for information bits and $\mathbf{B_P}$ corresponds to parity-check coefficients for parity bits, they are both $m \times m$ matrices. The first and last elements in the first column of $\mathbf{B_P}$ are "1", the second element is "0" and others are "−1", the rest of $\mathbf{B_P}$ is a $m \times (m-1)$ matrix with elements in its two diagonal lines all being "0" and the rest elements all being "−1".

2.2.2 Encoding Scheme

We denote information bits as $\boldsymbol{u} = \{\boldsymbol{u_1}, \boldsymbol{u_2},\ldots, \boldsymbol{u_m}\}$ and parity bits as $\boldsymbol{v} = \{\boldsymbol{v_1}, \boldsymbol{v_2},\ldots, \boldsymbol{v_m}\}$, the length of each $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$ is z, and assume that the sub-matrix $\mathbf{H_{i,j}}$ in the parity-check matrix corresponds to $p_{i,j}$ in the basis matrix. The encoding process based on parity-check matrix is accomplished as follows

1. Calculate intermediate variables $\boldsymbol{c_i}(i = 1, 2, \ldots, m)$ and then $\boldsymbol{v_1}$ afterwards

$$\boldsymbol{c_i} = \left(\sum_{j=1}^{m_b} \mathbf{H_{i,j}} \cdot \boldsymbol{u_j^T}\right)^T \tag{2.3}$$

$$\boldsymbol{v_1} = \sum_{i=1}^{m_b} \boldsymbol{c_i}. \tag{2.4}$$

2. Update intermediate variables $\boldsymbol{c_i}$ (just for $i = 1, 2$)

$$\boldsymbol{c_i} = \boldsymbol{c_i} + \left(\mathbf{H_{i,m+1}} \cdot \boldsymbol{v_1^T}\right)^T. \tag{2.5}$$

3. Calculate all other parity bits

$$\boldsymbol{v_i} = \begin{cases} \boldsymbol{c_1} & i = 2 \\ \boldsymbol{c_{i-1}} + \boldsymbol{v_{i-1}} & i \in [3,\ m]. \end{cases} \tag{2.6}$$

The encoded codeword is $\{\boldsymbol{u}, \boldsymbol{v}\} = \{\boldsymbol{u_1}, \boldsymbol{u_2},\ldots, \boldsymbol{u_m}, \boldsymbol{v_1}, \boldsymbol{v_2},\ldots, \boldsymbol{v_m}\}$ according to above steps, and note that all the additions referred are Modulo 2 additions.

2.3 Construction of Parity-Check Matrix

For most cases decoding algorithms for LDPC code are iterative, the information received by the nodes in the Tanner graph tends to be no more independent thus degrades the decoding performance when cycles exist in the Tanner graph or parity-check matrix. On the other hand, the existence of cycles can benefit the codewords by improving the minimum code distance, for which we should take both factors into consideration in the process of LDPC code design. Generally speaking, 4-cycle plays a negative role for the convergence of decoding process and considered having the worst influence on the performance of LDPC code and that we can achieve fairly good performance by removing it. The method given in this chapter absolutely avoids the existence of 4-cycle.

We will discuss the construction of the left part of parity-check matrix, i.e., what $\mathbf{B_S}$ corresponds since the right part $\mathbf{B_P}$ has been certain. Binary sparse sequence is a sequence with certain number of "1" and the others are all "0", the cross-correlation of two sequences s_1 and s_2 with the length of N is defined as

$$R_{12}(k) = \sum_{i=1}^{N} s_1(i) \cdot s_2((i+k) \bmod N). \tag{2.7}$$

No 4-cycle exists in the parity-check matrix means the cross-correlation value is less than 2 between any two arbitrary rows if we view each row as a sequence.

For each element in the m*z sparse sequence we label it with a number (0, 1,…, $m-1$) which equals to its subscript modulo m and be sure that there is one and only one "1" for all the elements that share a common label, thus there are m "1" in the entire sequence. Z sequences can be derived if we left cyclic shift the mentioned sequence by $k*m$ steps ($k = 1, 2,\ldots, z-1$), which can be used to constitute a matrix with size of $z \times (m*z)$. It's easy to find out that cross-correlation value $R_{12}(0)$ between any two arbitrary rows is 0 and quasi-cyclic form can be achieved by means of column exchange.

We should generate m sparse sequences which meet above requirements to get the parity-check matrix. The part which $\mathbf{B_P}$ corresponds to also has an effect on the cross-correlation value, that's why we also take it into consideration. The construction process is as follows:

1. First construct a matrix $\mathbf{A} = \text{zeros}(m, m*z)$ according to $\mathbf{B_P}$, $\mathbf{A}(i, j) = 1$ if $\mathbf{B_P}(i, j) == 0$ and we also set $\mathbf{A}(1, m*(z-1)+1) = 1$ and $\mathbf{A}(m, m*(z-1)+1) = 1$ which correspond to the two "1"s in $\mathbf{B_P}$.
2. Find m sparse sequences which should meet the requirements mentioned as Fig. 2.1 shows, Lcirshift (a, b) means left cyclic shift sequence a by length of b.
3. Obtain z sequences with each sparse sequence by left cyclic shift by $k*m$ steps ($k = 0, 1, 2,\ldots, z-1$) which also means a $z \times (m*z)$ matrix can be constructed with one sparse sequence if each row corresponds to a cyclic shifted sequence. A $(m*z) \times (m*z)$ matrix can be constructed with m sparse sequences found in step 2, left part of parity-check matrix $\mathbf{H_S}$ can be derived by its column exchange.
4. Concatenate $\mathbf{H_P}$ with $\mathbf{H_P}$, we get $\mathbf{H} = [\mathbf{H_S}|\mathbf{H_P}]$.

The removal of 4-cycle is ensured by limiting the cross-correlation values and sequences can still be found quickly with computer when m is large since the removal of 6-cycle is not considered. This method results in non-existence of "-1" in the basis matrix which makes the row weight and column weight be both m, the density is not low enough. We can set some of sub-matrices in $\mathbf{H_S}$ as zero matrix to solve this problem which has been verified to be an efficient solution.

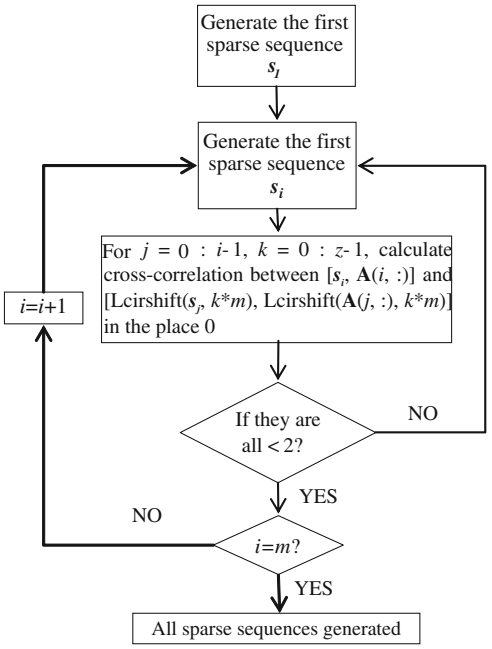


Fig. 2.1 Flowchart of generating all sparse sequences

2.4 Simulation Results

We set codeword length to 1,200 bits and coding efficiency 1/2 which is the same as one of L1C scheme, size of sub-matrix is set to 60 × 60 which means $z = 60$ and $m = 10$. The final parity-check matrix is obtained by set 60 sub-matrices in the left part of $\mathbf{H_S}$ to zero matrices randomly, (2.8) gives out the corresponding basis matrix as an example.

$$\mathbf{B} = \begin{bmatrix} -1 & -1 & 9 & 25 & -1 & -1 & -1 & 51 & 20 & -1 & 1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 16 & -1 & -1 & -1 & 50 & 55 & -1 & 31 & 41 & -1 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 26 & -1 & -1 & -1 & 12 & -1 & -1 & -1 & 7 & -1 & -1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 12 & -1 & 43 & -1 & -1 & 31 & -1 & 15 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & -1 & -1 & -1 \\ 41 & -1 & -1 & -1 & -1 & -1 & 3 & 28 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & -1 & -1 \\ 40 & -1 & -1 & 43 & 38 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 \\ -1 & -1 & 25 & 37 & 29 & -1 & -1 & -1 & 34 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 \\ -1 & -1 & -1 & -1 & 11 & -1 & -1 & 37 & 17 & 38 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & -1 \\ -1 & 56 & 13 & -1 & -1 & 5 & -1 & 47 & 33 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 \\ -1 & 28 & 42 & -1 & -1 & 6 & 23 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 \end{bmatrix} \tag{2.8}$$

LLR-BP algorithm is used as the decoding method to verify the performance of constructed LDPC code. Figure 2.2 gives out an example showing how bits error number changes with iteration times when SNR equals 1 dB which turns convergence successfully after 12 iteration steps, the parity-check matrix is constructed by the method as last chapter describes. This curve shows the contribution FEC makes for GNSS visually.

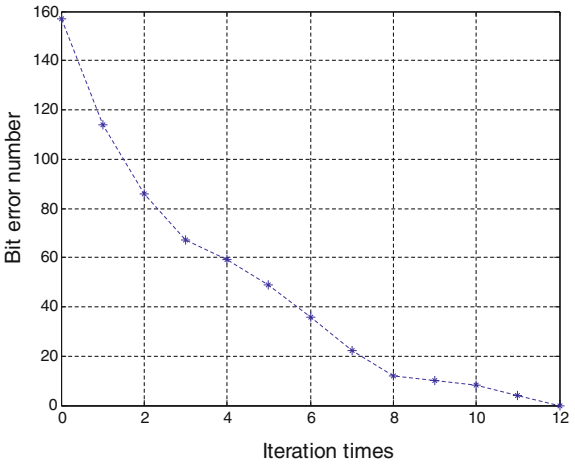


Fig. 2.2 An example showing relationship between bit error number and iteration times using LLR-BP decoding algorithm

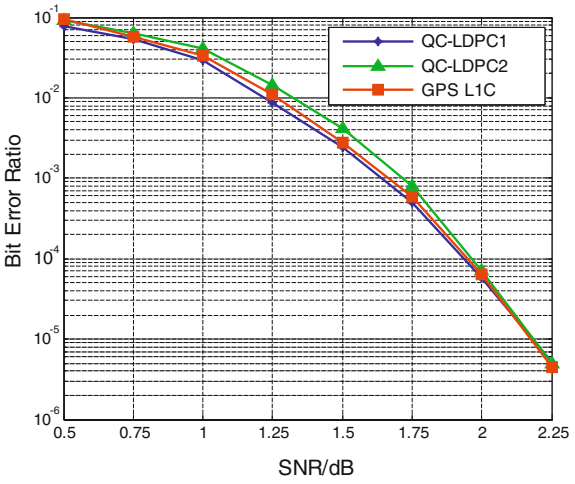


Fig. 2.3 BER performance comparison between QC-LDPC codes and GPS L1C LDPC codes

We simulated the bit error ratio (BER) performance of LDPC code used by GPS L1C signal, an optimized result [5] gives and our result for comparison to verify its performance. BPSK modulation is assumed to be used, i.e., "1" is mapped into "−1" and "0" is mapped into "1", and the channel is assumed to be AWGN. The maximum iteration times is set to 50 for each decoding process.

Figure 2.3 presents their BER performance curves, it's clearly shown that our result (blue curve) outperforms what [5] gives (green curve) and is slightly better than GPS L1C scheme when SNR is low, so it can be considered at least meet the performance of GPS L1C. Storage requirements are significantly reduced in the process of encoding and decoding with QC-LDPC scheme while achieving not lower performance.

2.5 Concluding Remarks

Considering the power of data channel is reduced by at least 3 dB due to the fact that generally half or more power of modernized GNSS signal is allocated to pilot channel, we can choose forward error correction (FEC) coding to make up for the loss of robustness of data demodulation due to power allocation. This paper briefly introduces the encoding scheme based on QC-LDPC code followed by proposing a method for constructing parity-check matrix without 4-cycle. Simulation result shows that its decoding performance can be as good as what GPS L1C uses. Further research can be focused on studying specific algorithm to "lighten" the check matrix which was randomly done for this paper.

References

1. Hu X, Tang Z et al (2009) Analysis on design principles of GPS and Galileo signal structure. Syst Eng Electron 31(10):2285–2293
2. GPS Navstar Joint Program Office. Navstar GPS space segment/user segment L1C interfaces [S], Draft IS-GPS-800, 2006
3. Richardson TJ, Urbanke RL (2001) Efficient encoding of low-density parity-check codes. IEEE Trans Inform Theory 47(2):P638–P656
4. Shanbao H, Liu CH, Yiming L (2009) Application of LDPC codes in satellite navigation systems. Spacecraft Eng 18(3):72–76
5. Qian H, Li GX, Chang J (2011) Application of quasi-cycle low-density parity check codes with high performance to satellite navigation signals. J Comput Appl 31(4):1145–1147
6. Wang Y, Yedidia JS, Draper SC (2008) Construction of high-girth QC-LDPC codes. 2008 5th International Symposium on Turbo Codes and Related Topics, Lausanne, 2008, pp 181–185
7. Li Z, Chen L, Zeng L et al (2006) Efficient encoding of qusai-cyclic low-density parity-check codes. IEEE Trans Comm 54(1):P71–P81

Chapter 3
ACED Multiplexing and Its Application on BeiDou B2 Band

Zheng Yao and Mingquan Lu

Abstract Existing dual-frequency constant envelope multiplexing (CEM) techniques, represented by alternate binary offset carrier (AltBOC) modulation, have strict constraint on component relative power distribution. Although it is desired to allocate more power to pilot channel to enhance measurement accuracy and tracking robustness, AltBOC is not able to meet this requirement. In this paper, a novel dual-frequency CEM technique is presented, in which with the total transmitted power fixed, different services in each sideband can have different power distribution, thus more power can be allocated on pilot channels to improve the tracking performance. The results are practical and provide a more internationally competitive solution for BeiDou navigation satellite system B2 signal.

Keywords Constant envelope multiplexing · Dual-frequency · Unequal power allocation · Navigation signal design

3.1 Introduction

In BeiDou satellite navigation system B2 band there is a high demand to the dual-frequency multiplexing. In this band, BeiDou plans to broadcast two QPSK signals located at B2a (1,207.14 MHz) and B2b (1,176.45 MHz) respectively, which carry different services [1]. Four-component alternate binary offset carrier (AltBOC) [2] is a feasible solution. And a combination of time division multiplexing and two-component AltBOC, named TD-AltBOC [3] is also proposed for B2.

When considering AltBOC and TD-AltBOC, they have obvious limitation. In AltBOC, components involved in multiplexed signal have to have equal power. Its

Z. Yao (✉) · M. Lu
Department of Electronic Engineering, Tsinghua University, Beijing, China
e-mail: yaozheng@ieee.org; yaozheng@tsinghua.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_3,

time-division form, i.e. TD-AltBOC, not only reduces the effective code length, which worsen the correlation characteristic between different PN codes, as discussed in [4], but also forces receivers to change correlator structure. In order to process it, receivers have to add time-division switching circuits. As analyzed in [5], the further augmentation of signals generated on the basis of time division multiplex in future is impossible.

Actually, an ideal multiplexing technique should be flexible enough to power allocation and chip waveform of components, so that the optimization of spreading modulation and power allocation do not have to compromise with multiplexing excessively. Moreover, multiplexing should be as *transparent* as possible to receivers. That means users do not need to know which multiplexing technique is being used in transmitter, and receivers will not be affected when this multiplexing technique is supplanted by more advanced scheme in future. However, the optimization potential is somewhat affected by the poor flexibility of existing techniques. More importantly, in the signal structure in modernized GNSS, it is desired to allocate more power to pilot channel to enhance measurement accuracy and tracking robustness, just like in GPS L1C [6]. However, AltBOC and its variant apparently do not possess such property.

This paper presents a new dual-frequency CEM technique, named asymmetric constant envelope double-sideband (ACED) multiplexing. It can combine no more than four signals with non-equal power at two different frequencies into a spectrum-split signal in which each sideband comprises no more than two codes modulated onto the orthogonal components. The power ratio between those channels can be arbitrary. Analysis shows that if BeiDou B2 signal employs this technique, it can provide better ranging performance than existing techniques.

3.2 ACED Multiplexing

Considering four different baseband spreading signals g_1, g_2, g_3, g_4, the spreading sequences of which are orthogonal to each other, without loss of generality, each channel can be represented as

$$s_i(t) = A_i g_i(t) d_i(t) \tag{3.1}$$

where d_i is the navigation data modulated in s_i, and A_i is the amplitude of channel i. Equality of those A_i is not required. However, for convenience of discussion, we normalize the total power of channels, that is $\sum_i A_i^2 = 1$.

When considering combining these four channels into an integrated signal in which s_1 and s_2 are located on the upper sideband, s_3 and s_4 are located on the lower sideband, and on each sideband two channels are modulated onto the orthogonal phases of the carrier, with the center frequencies of two sidebands $2f_{sc}$ apart, single-band modulation technique can be employed. However, if sine wave

complex subcarrier is employed directly, the envelope of the integrated baseband complex signal cannot be a const.

In ACED multiplexing, a novel 2-level waveform is employed as the complex subcarrier. The ACED multiplexing signal can be expressed as

$$\tilde{s}_l(t) = \frac{\sqrt{2}}{2}\alpha_1 \mathrm{sgn}[\sin(2\pi f_{sc}t + \varphi_1)] + \mathrm{j}\frac{\sqrt{2}}{2}\alpha_2 \mathrm{sgn}[\sin(2\pi f_{sc}t + \varphi_2)] \tag{3.2}$$

where

$$\begin{cases} \alpha_1 = -\sqrt{(s_1 + s_3)^2 + (s_2 - s_4)^2} \\ \alpha_2 = \sqrt{(s_1 - s_3)^2 + (s_2 + s_4)^2} \end{cases} \tag{3.3}$$

$$\begin{cases} \varphi_1 = -\ \mathrm{atan2}[(s_1 + s_3), (s_2 - s_4)] \\ \varphi_2 = \ \mathrm{atan2}[(s_2 + s_4), (s_1 - s_3)] \end{cases} \tag{3.4}$$

and atan2 $(\cdot, \cdot)$ is four-quadrant arctangent.

It can be verified that $\tilde{s}_l(t)$ has a constant envelope. More importantly, no power ratio limitation and constraint is imposed to the proposed technique, which implies more proportion of the transmission power can be allocated to pilot channels. As we show later, with this feature, the ranging measurement performance can be significantly improved.

Another notable advantage of the proposed technique is that in this multiplexing signal, any channel can be halted by simply set $A_i = 0$, while the constant envelope property remains unchanged. Therefore, in fact ACED technique can multiplex no larger than four signal channels, and during the running of the system, any signal channel can be selectively deactivated without affecting the other channels.

3.3 Applications in BeiDou B2 band

In the design of BeiDou satellite navigation system B2 band signals, to counter the shortage of AltBOC and its TD form, an improved scheme, in which ACED technique is used to realize the constant envelope combination of B2a and B2b is proposed in this section. In B2a and B2b there are two QPSK signals, in which the in-phase channels are data channels and quadrature-phase channels are pilot channels. So far there is no definite demand for allocating more power on one sideband than the other sideband (Actually, if really necessary, it is also easy to be satisfied by employing the proposed ACED technique). However, it is evidently true that more power should be allocated to the pilot channels, if possible, to improve ranging performance.

In the recommendation, the power ratio of pilot channel and data channel in each sideband are all 3, and the upper and lower sidebands have the equal total power, i.e., $A_1 : A_2 : A_3 : A_4 = 1 : \sqrt{3} : 1 : \sqrt{3}$. This signal is denoted as ACED(3) signal hereafter. In frequency domain, Fig. 3.1 depicts the sketch of the in-phase and quadrature-phase components of this signal.

From (3.3) and (3.4) it is easy to verify that in the modulation constellation, this signal has 12 phase points which are equispaced on the unit circle. That is, the integrated signal can be described as a 12-PSK signal

$$\tilde{s}_l(t) = \exp\left(\mathrm{j}\frac{\pi}{6}x\right) \tag{3.5}$$

with x ranging from 1 to 12. The modulation constellation and the index of each phase point are shown in Fig. 3.2. For each value combination

$$\boldsymbol{v}_k \triangleq (g_1 d_1, g_2 d_2, g_3 d_3, g_4 d_4)^{\mathrm{T}} \tag{3.6}$$

$\tilde{s}_l(t; \boldsymbol{v}_k)$ has oscillation at f_{sc} and vibrates between two phase points which are symmetrical about the origin point of the constellation. Within a single subcarrier period $t \in [nT_{sc}, (n+1)T_{sc})$, different $\boldsymbol{v}_k$ has different switching time point. However, those time points are equispaced in the interval $[nT_{sc}, (n+1)T_{sc})$. In the transmitter, the mapping table between $\boldsymbol{v}_k$ and x is presented in Table 3.1.

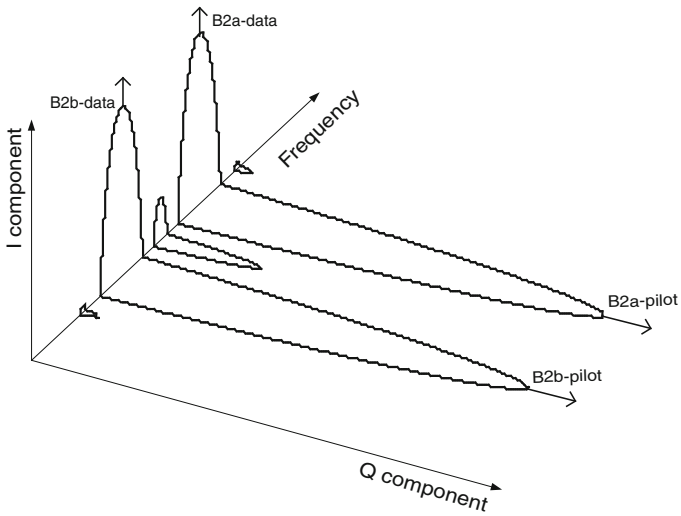


Fig. 3.1 Spectra of proposed ACED(3) signal

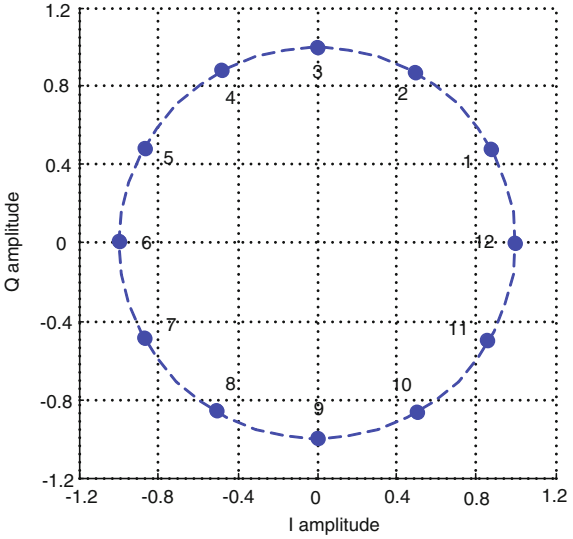


Fig. 3.2 Constellations of ACED(3) signal

3.4 Performance Analysis

3.4.1 Spectrum Characteristic

With the assumption of ideal auto-correlation and cross-correlation characteristics of spreading sequence employed by $s_i(t)$, after tedious mathematical derivation one can obtain the power spectral density (PSD) expression of $s_i(t)$

$$G(f) = \frac{f_c \cos^2\left(\frac{\pi f}{f_c}\right)\left\{1 - \cos(6\psi)[\sin(5\psi)\sin\psi + \cos^2\psi]\right\}}{2\pi^2 f^2 \cos^2(6\psi)} \tag{3.7}$$

with $\psi \triangleq \frac{\pi f}{12 f_{sc}}$.

For B2 case, PSD of ACED(3) is depicted in Fig. 3.3. For comparison, PSDs of AltBOC and its time-division form are also drawn in the same figure.

In PSD, only the sum of the power in in-phase component and quadrature-phase component can be reflected. Since in the given restrained power ratio condition, the total power in upper sideband and lower sideband are equal, in Fig. 3.3 it can be seen that the PSD of ACED(3) is symmetrical. And it is interesting to note that its spectrum shape is very similar to AltBOC and its time-division form.

Actually, by comparing (3.7) with AltBOC one can also find that ACED(3) differs from AltBOC only in the high frequency terms which uses the sum of third-harmonic term and sixth-harmonic term instead of fourth-harmonic term. Consequently, when ACED signal is treated as a whole, for those receivers with a double-side bandwidth less than 100 MHz, the assumption of which can cover all the practical receivers, it does not have significant performance difference to AltBOC and its time-division form.

Table 3.1 Look up table for ACED(3) phase states

$g_1 d_1$	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1
$g_3 d_3$	1	1	1	1	−1	−1	−1	−1	1	1	1	1	−1	−1	−1	−1
$g_2 d_2$	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1	1	1	−1	−1
$g_4 d_4$	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1
t mod T_{sc}	x according to $\tilde{s}_l(t;\boldsymbol{v}_k) = \exp(\mathrm{j}\, x\pi/6)$															
$\left[0, \frac{1}{12}\right)T_{sc}$	2	12	12	10	3	5	1	9	3	7	11	9	4	6	6	8
$\left[\frac{1}{12}, \frac{2}{12}\right)T_{sc}$	2	6	12	10	3	5	1	9	3	7	11	9	4	6	12	8
$\left[\frac{2}{12}, \frac{3}{12}\right)T_{sc}$	2	6	12	10	3	5	1	3	9	7	11	9	4	6	12	8
$\left[\frac{3}{12}, \frac{4}{12}\right)T_{sc}$	8	6	12	4	3	5	1	3	9	7	11	9	10	6	12	2
$\left[\frac{4}{12}, \frac{5}{12}\right)T_{sc}$	8	6	12	4	9	5	1	3	9	7	11	3	10	6	12	2
$\left[\frac{5}{12}, \frac{6}{12}\right)T_{sc}$	8	6	6	4	9	5	1	3	9	7	11	3	10	12	12	2
$\left[\frac{6}{12}, \frac{7}{12}\right)T_{sc}$	8	6	6	4	9	11	7	3	9	1	5	3	10	12	12	2
$\left[\frac{7}{12}, \frac{8}{12}\right)T_{sc}$	8	12	6	4	9	11	7	3	9	1	5	3	10	12	6	2
$\left[\frac{8}{12}, \frac{9}{12}\right)T_{sc}$	8	12	6	4	9	11	7	9	3	1	5	3	10	12	6	2
$\left[\frac{9}{12}, \frac{10}{12}\right)T_{sc}$	2	12	6	10	9	11	7	9	3	1	5	3	4	12	6	8
$\left[\frac{10}{12}, \frac{11}{12}\right)T_{sc}$	2	12	6	10	3	11	7	9	3	1	5	9	4	12	6	8
$\left[\frac{11}{12}, 1\right)T_{sc}$	2	12	12	10	3	11	7	9	3	1	5	9	4	6	6	8

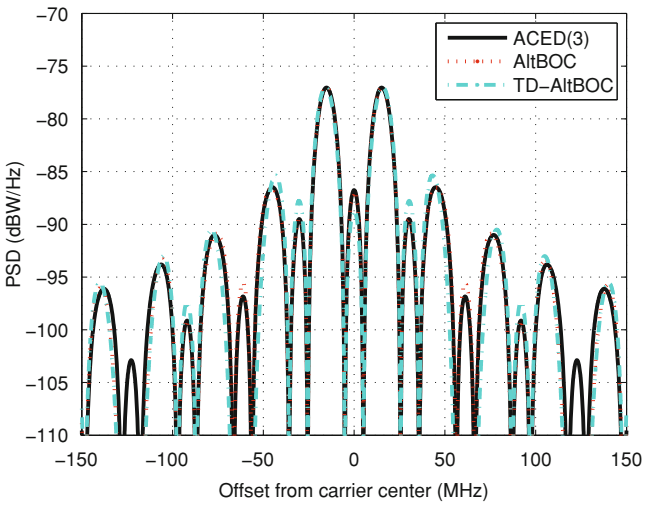


Fig. 3.3 Power spectral density of ACED(3), AltBOC and TD-AltBOC

Although theorists keen to analyze and compare the theoretical performance under full-spectrum matching processing with an over-wide receiving bandwidth, due to the different and unknown navigation bits modulated on two of the signal components, under most conditions, full-spectrum matching processing does not have operable. Moreover, the over-wide receiving bandwidth required by full-

spectrum matching processing in front-end and the high sample rate constraint it brings are out of reach for ordinary receivers. In next section, we focus our analysis on performance under a more practical separate processing mode, in which the validity of power relationship design constraint among four components can be also verified.

3.4.2 Power Constraints and Correlation Loss

Though not just AltBOC, in theory any constant envelope multiplexing integrated signal has the feasibility of full-spectrum matching processing, for normal requirement, constant envelope multiplexing ought to be transparent to receivers. That means users does not need to know which constant envelope multiplexing technique is being employed in transmitter. Receivers only need to generate the replica of the desired component and calculate its correlation with the received signal.

The correlation power loss η is defined as the ratio between the sum of all component signals power measured at the correlator output and the power of $\tilde{s}_l(t)$ in decibel. Since the total power of $\tilde{s}_l(t)$ has been normalized, the correlation power loss can be expressed as

$$\begin{aligned}\eta &= \sum_{i=1}^{4}|R_i|^2 = \sum_{i=1}^{4}\left|\frac{1}{T}\overline{\int_T \tilde{s}_l(t)\hat{s}_i(t)\mathrm{d}t}\right|^2 \\ &= \frac{\sum_{i=1}^{4}\left\{\sum_{\ell\in\Xi_i}\int_0^{T_{sc}}\tilde{s}_l(t;\boldsymbol{v}_\ell)e^{-\mathrm{j}2\pi f_{sc}t}\mathrm{d}t - \sum_{\ell\in\Xi_i^{\complement}}\int_0^{T_{sc}}\tilde{s}_l(t;\boldsymbol{v}_\ell)e^{-\mathrm{j}2\pi f_{sc}t}\mathrm{d}t\right\}^2}{256T_{sc}^2}\end{aligned} \tag{3.8}$$

where $\Xi_i = \{\ell | g_i = 1 \text{ in } \boldsymbol{v}_\ell\}$, and R_i is the average output of the correlator for s_i.

Referring to the phase states of each $\boldsymbol{v}_\ell$ within $[0, T_{sc})$, which are shown in Table 3.1, and employing (3.8), one can obtain that the average correlator output power of four components are

$$\begin{cases}|R_1|^2 = |R_3|^2 = 0.1013 \\ |R_2|^2 = |R_4|^2 = 0.3040\end{cases} \tag{3.9}$$

the ratio of which is identical to the predefined value. And the normalized correlation power loss is $\eta = 0.91$ dB, which exactly corresponds to the fraction of fundamental harmonic in square wave complex sub-carrier. As analyzed in [2], this means that ACED does not reduce the desired signal power in main lobes.

It should be noted that, the definition correlation power loss here has some distinction with the multiplexing efficiency. Correlation power loss is under the assumption of separately handling each channel as BPSK signal, while the so-called AltBOC multiplexing efficiency 85.36 % is measured under the

assumption of unlimited bandwidth and complex 8-level local waveform as given in [7]. As analyzed in [2] and validated in next section, for BPSK-like receiving with a finite bandwidth, the correlation power loss of both AltBOC and TD-AltBOC are also 0.91 dB, as same as that of ACED(3).

3.4.3 Receiving Performance

By using the similar procedure of (3.8), one can calculate the correlator output power of four components when each channel of AltBOC is handled as BPSK signal while each channel of TD-AltBOC is handled as TD-BPSK signal. Under the assumption that the total power of integrated signals are normalized, for both AltBOC and TD-AltBOC, they have

$$|R_1|^2 = |R_2|^2 = |R_3|^2 = |R_4|^2 = 0.2026 \tag{3.10}$$

Consequently, when signals are processed as four BPSK channel, the correlation power losses are identical, which is 0.91 dB. This means with a BPSK-like receiver, though different dual-frequency multiplexing signals may have different nominal multiplexing efficiencies, the actual total power in correlators are the same. Nevertheless, thanks to the higher power allocation of pilot channels in ACED(3), with the same tracking loop structure, ACED(3) can have higher accuracy and higher sensitivity in both carrier tracking and code tracking.

Figure 3.4 shows the pilot channel theoretical code tracking error of ACED(3) and TD-AltBOC, with the same equivalent noise bandwidth $B_L = 5$ Hz, 20 ms coherent integration time, and 1 chip spacing between early and late correlators. It can be seen from the figure, because of the greater power weight on pilot channel, comparing with AltBOC and TD-AltBOC, ACED(3) has 1.8 dB improvement in tracking sensitivity threshold, which also means, with the same total transmission power, the ranging error variance of ACED(3) introduced by thermal noise is only $2/3$ that of TD-AltBOC. The error variance ratio between ACED(3) and TD-AltBOC is shown in Fig. 3.5, which can also represent the better tracking performance of ACED(3). The outstanding ranging performance can give BeiDou B2 signals superiority over the other GNSS in the same band in future navigation applications.

Although in ACED(3), the power allocation of data channel is less than that in AltBOC, as analyzed in [6], by using low data rate and advanced channel encoding technique, the decrease of demodulation performance can be within acceptable levels. One should note that both GPS L1C and BeiDou B1C use 1 : 3 power allocation scheme, which has become a trend in next generation signal design.

When considering interoperability, ACED(3) signal of which the central frequency is at 1,191.795 MHz, can be treated as two QPSK signals located at 1,176.45 and 1,207.14 MHz respectively and can share antenna and radio-frequency front-end with Galileo E5a, E5b, and GPS L5 signals. Moreover, the

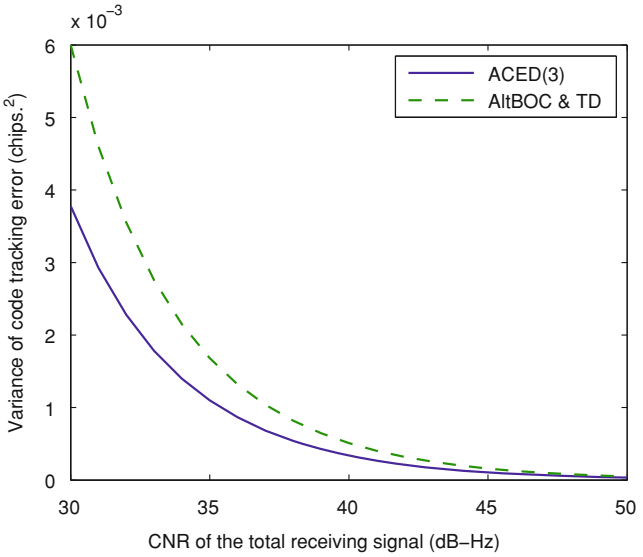


Fig. 3.4 Variance of code tracking errors for ACED(3) and (TD)AltBOC

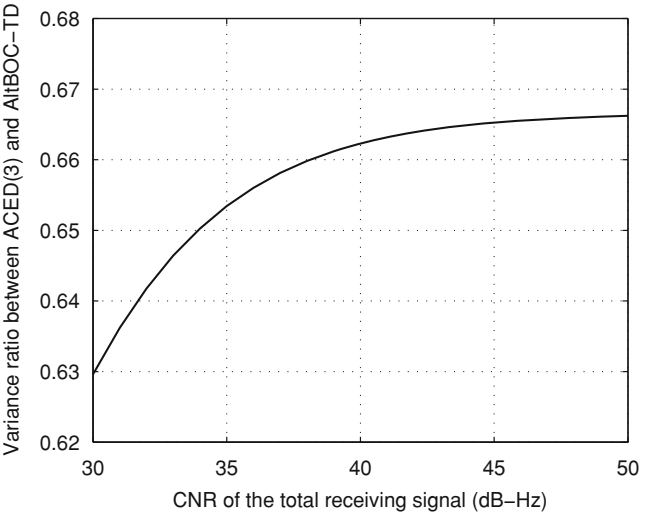


Fig. 3.5 Tracking performance improvement of ACED(3) comparing with (TD)AltBOC

processing channels of the signals above in multi-system receivers can have exactly the same structure without having to add additional time division switching circuit. At this point, ACED(3) is superior to TD-AltBOC once again.

3.5 Conclusions

For BeiDou satellite navigation system B2 band signals, to counter the shortage of existing AltBOC and its TD form, this paper gives an alternative based on the proposed technique. Analysis shows that under the wideband full-spectrum matching processing, this signal has a quite similar performance to AltBOC and TD-AltBOC, while in BPSK-like receiving mode, comparing with the existing B2 signal, it provides a significant performance improvement, which can give BeiDou B2 signals superiority over the other GNSS in the same band in future navigation applications. Consequently, this technique can serve as a practical and competitive alternative to the existing B2 signal scheme.

Acknowledgments This work is supported by the National Science Foundation of China (NSFC), Grant 61201190.

References

1. ICG (2010) Current and planned global and regional navigation satellite systems and satellite-based augmentations systems. In: International committee on global navigation satellite systems providers' forum, New York
2. Lestarquit L, Artaud G, Issler J-L (2008) AltBOC for dummies or everything you always wanted to know about AltBOC. In: ION GNSS, 2008 Savannah, GA, US, pp 961–970
3. Tang Z, Zhou H, Wei J et al (2010) TD-AltBOC: a new COMPASS B2 modulation. In: China satellite navigation conference 2010, Shanghai, China
4. Betz J, Cahn CR, Dafesh PA et al (2006) L1C signal design options. In: Technical national meeting of the institute of navigation. CA, Monterey, pp 685–697
5. Kharisov V, Povalyaev A (2011) Optimal aligning of GNSS navigation signals sum. In: The 24th international technical meeting of the satellite division of the institute of navigation. Portland, Oregon, pp 3141–3155
6. Betz JW, Blanco MA, Cahn CR et al (2006) Description of the L1C signal. In: ION GNSS 19th international technical meeting of the satellite division. Fort Worth, TX, US, pp 2080–2091
7. Rebeyrol E, Julien O, Macabiau C et al (2007) Galileo civil signal modulations. GPS Solut 11(3):159–171
8. Dafesh PA, Cahn CR (2011) Application of POCET method to combine GNSS signals at different carrier frequencies. In: International technical meeting of the institute of navigation. San Diego, CA, USA
9. Avila-Rodriguez JA (2008) On generalized signal waveforms for satellite navigation. University FAF Munich, Munich

Chapter 4
Spacecrafts Navigation Signal Research Based on GNSS Constellation

Peng Li, Zhonggui Chen, Yanan Gu and Yang Si

Abstract This paper analyzes the navigation signal design for space applications, especially for the high earth orbit spacecrafts which have a severely limited view of the main beams of the traditional GNSS signals. Based on the MEO constellation and antenna direction, a navigation signal is derived for space applications. This signal is farther analyzed and simulated according to the transmit antenna pattern, visibility, navigation capability and reception. The results show that the minimum received signal power is −167.3 dBW when the transmit power is 20 dBW and the transmit beam half angle is 60°. Simulation shows that the geostationary users can achieve the visibility of 11 satellites and the positioning precision of about 18.2, which can meet the navigation requirements of high earth orbit spacecraft. The analysis in this paper provides a possible solution for the space navigation signal design in the Chinese Compass construction.

Keywords Spacecraft navigation · Signal design

4.1 Introduction

GNSS is one of the essential national infrastructures, which provides navigation and positioning services for land users, ships and sky users. The demand for middle orbit or high orbit spacecraft navigation has been getting stronger and stronger as the national interest had extended to the outer space. The navigation satellites provide a series of services for the spacecraft, including orbit determination, attitude determination and time synchronization. The spacecraft navigation based on GNSS is becoming one of the research focuses in GNSS design area.

P. Li (✉) · Z. Chen · Y. Gu · Y. Si
China Academy of Space Technology, Beijing, People's Republic of China
e-mail: richardlee0914@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_4,

In 1982, the satellite "Landsat-4" is the first spacecraft which were equipped with GPS receivers and had get the navigation service [1]. From then on, the spacecraft navigation consideration is studied and GPS had been applied quickly in space navigation [2]. GPS is not only used for the spacecraft's orbit determination [3] but also used for its attitude determination [4]. A lot of experimental researches have been conducted on the application of high orbit spacecraft navigation. According to the experimental result of high orbit spacecraft GPS receiver, the "narrow beam" design of antenna pattern will lead to bad visibility when receive the main lobe signal and bad receiving caused by axial ratio degradation when receive the minor lobe signal [5, 6].

The navigation demand of low orbit users (height: 0–3,000 km) has been satisfied as the GNSS and its antenna pattern is design for the land users and near land users. The GDOP of GPS satellites will degraded as the height of the middle orbit users (height: 3,000–8,000 km) increase, which results in the degradation of the signal availability. When the high orbit users or geostationary users are concerned, only the signal outside the shadow of the earth can be received. As a result, the visibility will be reduced greatly besides GDOP degradation. The deep space users will suffer not only the problem of visibility but also the low signal power and high GDOP caused by the long distance. The main requirement for the high orbit users and geostationary users is that one of the GPS signal is received properly, which ensures the accurate timing without expensive atomic clock [7].

If MEO constellation is used to provide navigation service for high orbit and geostationary users, the difficulties of visibility and GDOP mentioned above will be encountered. Some paper presented that the multi-system compatibility improve the visibility [8]. But space navigation is related with the national security, the multi-system compatibility is limited in this area. Other researchers fix the antenna opposite the earth to serve the high orbit users [9]. However, this solution changes the satellite greatly, and is infeasible.

This paper proposes a spacecraft navigation signal, which is bases on the existing GNSS MEO constellation, and the antenna is fixed to the earth. Firstly, the visibility and GDOP is analyzed to calculate the main lobe. And then the transmit antenna pattern is designed. At last the link analysis and navigation accuracy is conducted on the spacecraft navigation signal.

4.2 Signal Beam Design for the Spacecraft Navigation Signal

The beam width design of spacecraft navigation signal is needed when the antenna is fixed to the earth. Visibility and GDOP is required in the beam width design.

4.2.1 Visibility Analysis

The visibility of navigation satellites is related with the beam width and the height of the users. Figure 4.1 shows the visible area of the users with different orbit.

As the antenna is fixed to the earth, the visible area of some users contained the area inside the signal beam. The users conclude low orbit user, middle orbit user and parts of the high orbit user, which is lower than the MEO navigation satellite. The users higher than the MEO navigation satellite must receive the signal outside the shadow of the earth.

Simulation is used to analyze the visibility of the users with different orbit height in different signal beam width. The simulation configuration is as follows:

1. Spacecraft navigation signal is based on MEO GNSS constellation, which is Walker 24/3/2, the orbit inclination is 55°, and the orbit height is 21,528 km.
2. The half angle of the available navigation signal area is 13.2° for the land users, and the half angle is 15.3° for the sky users 1,000 km above the earth.
3. The orbit height is respectively 3,000, 8,000, 20,000, 35,786 and 380,000 km for low orbit users, middle orbit users, high orbit users, geostationary users and deep space users.

The 1 week simulation is fulfilled with different user and different signal half angle. The time rate when the visibility is above 4 satellites is shown in Table 4.1

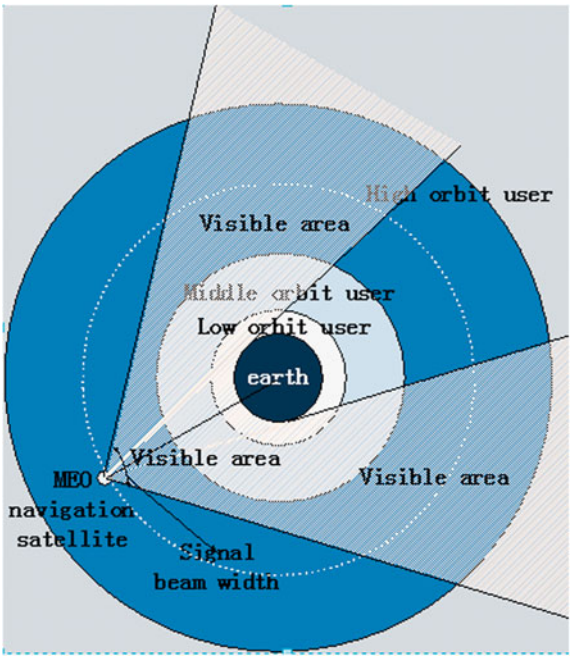


Fig. 4.1 Visibility for the satellite navigation signals

Table 4.1 The time rate when the visibility is above 4 satellites

Users type	20° (%)	30° (%)	40° (%)	50° (%)	60° (%)
Low orbit (3,000 km)	100	100	100	100	100
Middle orbit (8,000 km)	26.4	100	100	100	100
High orbit (20,000 km)	11.1	100	100	100	100
Geostationary(35,786 km)	0	39.78	100	100	100
Deep space (380,000 km)	0	0	19.4	92.9	100

The result illustrates 100 % low orbit users have visibility of above 4 satellites in all the half angle conditions. When the half angle is greater than 30°, the 4 satellites visibility can be certainly achieved for the middle orbit and high orbit users. And the geostationary user and deep space user need 40° and 60° half angle respectively to achieve 100 % 4 satellites visibility.

4.2.2 Geometric Dilution of Precision Analysis

The distribution of the satellites must meet the requirement of spacecraft navigation besides the visibility, which is measured by GDOP [10].

$$GDOP = \frac{\sqrt{\sigma_{x_u}^2 + \sigma_{y_u}^2 + \sigma_{z_u}^2 + \sigma_{ct_b}^2}}{\sigma_{UERE}} \tag{4.1}$$

In which, σ_{UERE} is use equivalent range error, $\sigma_{x_u}^2$, $\sigma_{y_u}^2$, $\sigma_{z_u}^2$ *and* $\sigma_{ct_b}^2$ are the three dimensions error and the time error respectively. GDOP represents the degradation from the range error to positioning error.

The following simulation analyzes the GDOP for different users in different beam width. Simulation configuration is unchanged as last section (Table 4.2).

The result shows:

1. The GDOP is small and constant for the low orbit users, because the low orbit users' activity area is covered by the signal beam when the half angle is greater than 20°.
2. The GDOP is small and constant for the middle orbit users, when the half angle is greater than 40°, and the users' activity area is covered by the signal beam

Table 4.2 Average GDOP for different space users

Users type	20°	30°	40°	50°	60°
Low orbit (3,000 km)	0.8932	0.8932	0.8932	0.8932	0.8932
Middle orbit (8,000 km)	NaN	0.9031	0.7695	0.7695	0.7695
High orbit (20,000 km)	NaN	18.48	6.724	3.399	1.979
Geostationary (35,786 km)	NaN	NaN	33.05	15.43	9.06
Deep space (380,000 km)	NaN	NaN	NaN	NaN	1697.6

3. The GDOP grows when the half angle is increased for the high orbit, geostationary and deep space users. The high orbit user's GDOP is less than 10 when the half angle is greater than 40°. The geostationary user's GDOP is less than 10 when the half angle is greater than 60°.
4. The deep space user's GDOP is unsatisfactory as the users is far from the MEO satellites.

4.2.3 Summary of the Signal Beam Design

The visibility and GDOP simulation analysis can be sum up as:

1. The existing GNSS signal narrow beam design can satisfy the requirement of low orbit users and parts of the middle orbit users.
2. The MEO based GNSS constellation can't provide the positioning service to the deep space users.
3. The spacecraft navigation can be achieved based on the existing GNSS constellation for the high orbit and geostationary users when the signal is design as wide beam.

4.3 Antenna Pattern Design

In order to get satisfactory visibility and GDOP of the high orbit navigation, the antenna beam width must be large enough. When the geostationary user is concerned, the spacecraft navigation geometric illustration is shown as Fig. 4.2.

The antenna beam half angle $\beta = 60°$, the earth shade angle $\gamma = 13.2°$, MEO satellite orbit radius $R_m = 27,878$ km, GEO orbit radius $R_g = 42164.2$ km, When the transmitting angle is α, the distance between navigation satellite and high orbit spacecraft R_d is:

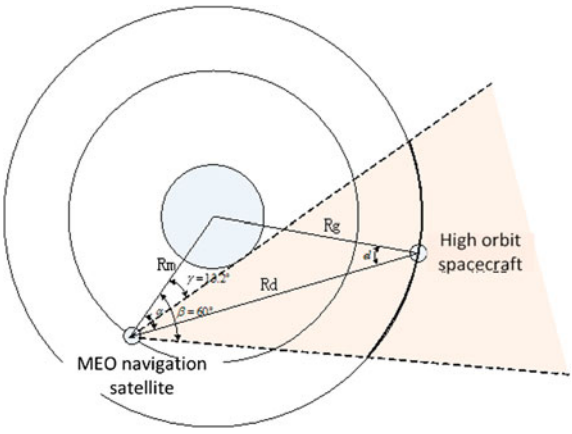


Fig. 4.2 Space navigation signal geometry illustration

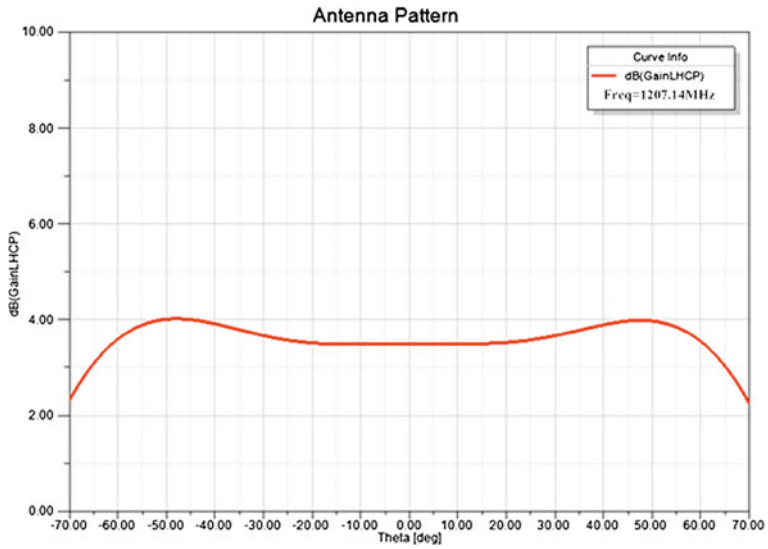


Fig. 4.3 Space navigation signal antenna pattern

$$R_d = R_m \cdot \cos\alpha + \sqrt{R_g^2 - (R_m \sin\alpha)^2} \tag{4.2}$$

The transmitting loss is in direct ratio to the square of R_d, as a result, the optimized design is that the transmitting antenna gain is also in direct ratio to the square of R_d in the range of 13.2° ~ 60°. The antenna pattern should be designed as the gain between 0 ~ 32° is reduced and increased between 13.2° ~ 60°.

According to the rule above, a small-sized backfire antenna is adopted to adjust the antenna gain. The gain is reduced between 0 ~ 13.2° of the antenna beam in order to assure the its requirement between 13.2° ~ 60°, and decline quickly outside 60°. Furthermore, the antenna has excellent polarization characteristics with lower minor lobe. As a result, the proposed antenna is capable of space applications.

The proposed antenna pattern is shown in Fig. 4.3. The signal beam gain is listed in Table 4.3. The transmitting gain is lower than the traditional "saddle shape" signal beam designed for the land users as the spacecraft navigation signal has wide beam, and the max-gain is in about ±50°.

4.4 Link Analysis

The formula of high orbit satellite received power is:

$$P_r = P_t + G_t - L_S - L_a + G_r - L_p \tag{4.3}$$

Table 4.3 Transmit antenna gain

Transmitting angle	Gain (dBi)	Transmitting angle	Gain (dBi)
0°	3.5	±40°	3.9
±5°	3.49	±45°	4.0
±10°	3.48	±50°	4.0
±15°	3.49	±55°	3.9
±20°	3.51	±60°	3.6
±25°	3.57	±65°	3.09
±30°	3.67	±70°	2.3
±35°	3.8		

where P_t is transmitting power, G_t is transmitting antenna gain, L_S is free space attenuation, L_a is atmospheric loss, G_r is receiving antenna gain, L_p is antenna mismatch loss. As the space navigation signals do not pass through the atmosphere, the atmospheric loss $L_a = 0dB$.

When the satellite antenna radiation angle α change from 15° to 60°, the high orbit spacecraft received signal inclination el change from 9.9° to 35°.

$$el = \sin^{-1}\left[\frac{R_m \sin(\alpha)}{R_g}\right] \tag{4.4}$$

The transmission distance of the signal space shall be calculated according to formula (4.2). When the transmit power is 20 dBW, the transmitting EIRP changed as antenna transmitting angle in the receiving area is shown in Fig. 4.4. When COMPASS receiver antenna gain is 0 dBi, the high orbit spacecraft received power of different heights changed as antenna transmitting angle in the receiving area is shown in Fig. 4.5, the link budget is shown in Table 4.4.

4.5 Navigation Accuracy Analysis

The accuracy of high orbit spacecraft navigation signal measured by the positioning error variance,

$$\sigma_u = GDOP \cdot \sigma_{UERE} \tag{4.5}$$

Where σ_{UERE} is the user equivalent range error. σ_{UERE} is calculated as follow:

$$\sigma_{UERE} = \sqrt{\sum_i \sigma_i^2} \tag{4.6}$$

In the formula, σ_i represents the receiver pseudorange measurement error, ephemeris error, satellite clock error and multipath effect etc. When the uncorrelated cumulative receiver is used, the receiver pseudorange measurement error can be calculated as [10],

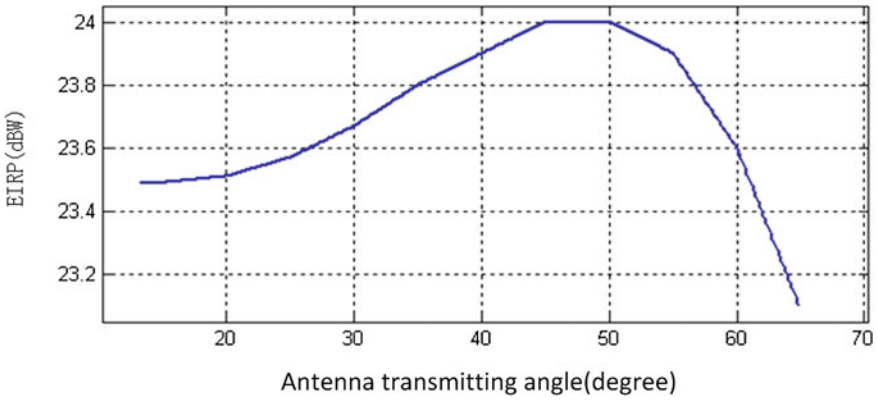


Fig. 4.4 Transmitting EIRP pattern

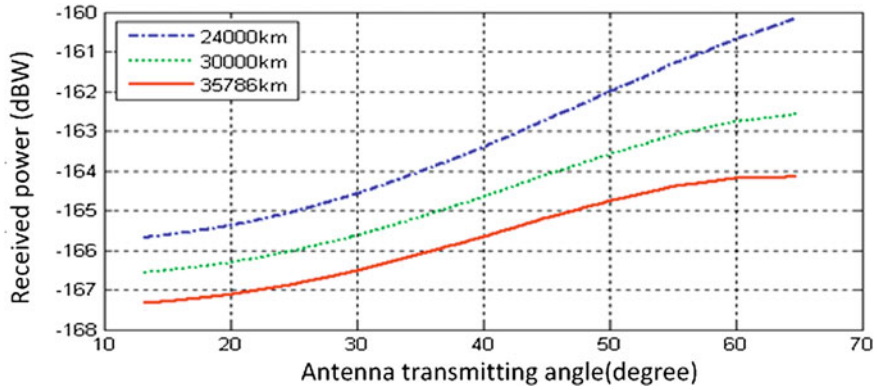


Fig. 4.5 Receiving power pattern

Table 4.4 High orbit navigation receiving power budget

Link parameter	24,000 km	30,000 km	35,786 km	Note
Transmitting power P_t (dBW)	20	20	20	
EIRP (dBW)	23.1	23.1	23.1	Min
Free space attenuation L_S (dB)	189.2	190	190.8	Max
Atmospheric loss L_a (dB)	0	0	0	
Gr (dBi)	0	0	0	
Received power P_r (dBW)	−165.7	−166.6	−167.3	Min

$$\sigma_1 = \sqrt{\frac{B_L(1-0.5B_LT)\int_{-\beta_r/2}^{\beta_r/2} G_w(f)G_s(f)\sin^2(\pi f\Delta)df}{(2\pi)^2 C_s\left(\int_{-\beta_r/2}^{\beta_r/2} fG_s(f)\sin(\pi f\Delta)df\right)^2}\cdot\left[1+\frac{\int_{-\beta_r/2}^{\beta_r/2} G_w(f)G_s(f)\cos^2(\pi f\Delta)df}{TC_s\left(\int_{-\beta_r/2}^{\beta_r/2} G_s(f)\cos(\pi f\Delta)df\right)^2}\right]} \tag{4.7}$$

where B_L is loop bandwidth, $G_w(f)$ is carrier power, $G_s(f)$ is signal power spectral density, T is the accumulated time. When the space navigation signal is BPSK (5) and the received power is as it is shown in Table 4.4, the equivalent range error budget of high orbit spacecraft for different orbit height is shown in Table 4.5.

4.6 High Orbit Navigation Simulation

4.6.1 Simulation Configuration

This section selects three sphere whose orbit heights are 24,000, 30,000 and 35,786 km, analysis the visibility and accuracy distribution of high orbit spacecraft in the three sphere.

1. The space navigation signal based on MEO global navigation system constellation, which use Walker 24/3/2 constellation configuration, with inclination of 55° and height of 21,528 km;
2. The navigation satellite transmitting antenna point to the ground, with the beam width of ±60°, transmitting antenna gain as shown in Table 4.3;
3. the received power of the three sphere surface and the UERE is calculated according to the minimum received power and UERE as shown in Tables 4.4 and 4.5;
4. The simulation of visibility on space navigation signal simulation use spherical polar coordinate grid of 1°, the time interval is half an hour.

Table 4.5 Space navigation signal UERE budget

Overall Error(1σ)	24,000 km (m)	30,000 km (m)	35,786 km (m)
Pseudorange measurement error σ_1	0.62	0.69	0.75
Ephemeris error σ_2	0.5	0.5	0.5
Satellite clock error σ_3	1.7	1.7	1.7
Multipath effect σ_4	0.1	0.1	0.1
Equivalent range error σ_{UERE}	1.80	1.87	1.93

4.6.2 Visibility Simulation

Figures 4.6, 4.7 and 4.8 illustrate the navigation signal visibility for the spacecraft in the sphere with radius of 24,000, 30,00 and 35,768 km respectively. The outline of the land is used to indicate the latitude and longitude, not because the receiver are placed on the land.

The simulation shows:

1. All the spacecraft receive the navigation signal outside the shadow of the earth, because the simulated sphere is higher than 21,528 km (height of the MEO navigation satellite). The visibility is degradated when the orbit is getting higher.

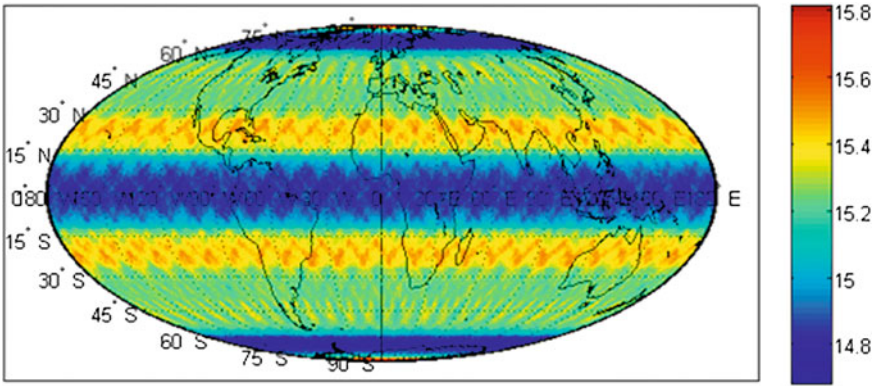


Fig. 4.6 The visible satellites number simulation for sphere with orbit of 24,000 km

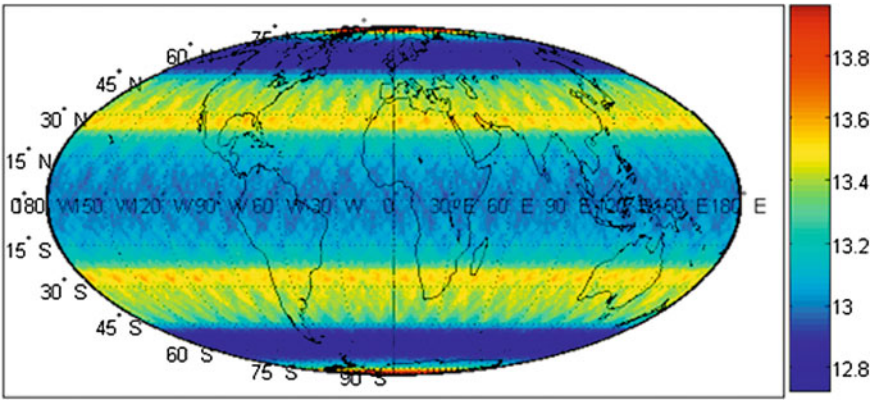


Fig. 4.7 The visible satellites number simulation for sphere with orbit of 30,000 km

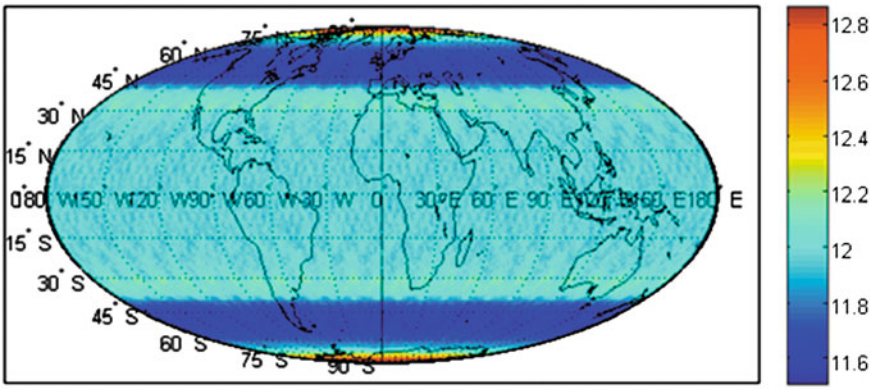


Fig. 4.8 The visible satellites number simulation for geostationary orbit sphere

2. The available satellite number is more than 14 in the sphere of 24,000 km, and the available satellite number is more than 11 in geostationary orbit sphere.
3. The available satellite number can achieve more than 11 between the area of MEO orbit sphere and geostationary orbit sphere, if the transmitting beam width of ±60° is adopted.

4.6.3 Navigation Accuracy Simulation

Figures 4.9, 4.10 and 4.11 illustrate the navigation accuracy for the spacecraft in the sphere with radius of 24,000, 3,000 and 35,786 km respectively.

The simulation shows:

1. The accuracy is degradated when the orbit is getting higher.

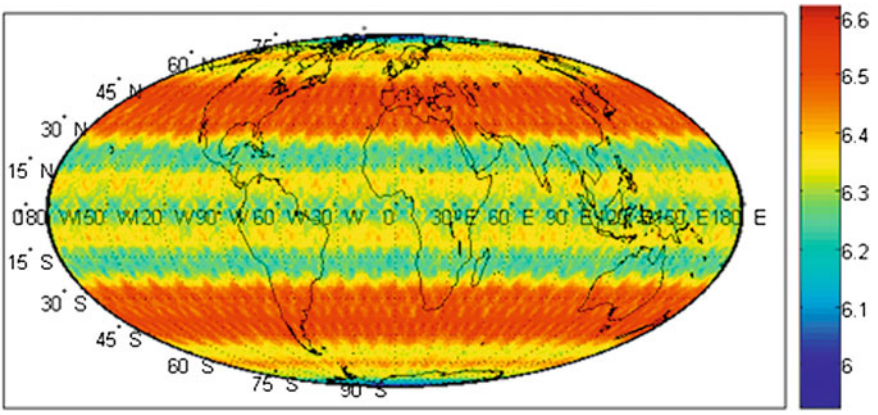


Fig. 4.9 Positioning precision simulation for sphere with orbit of 24,000 km

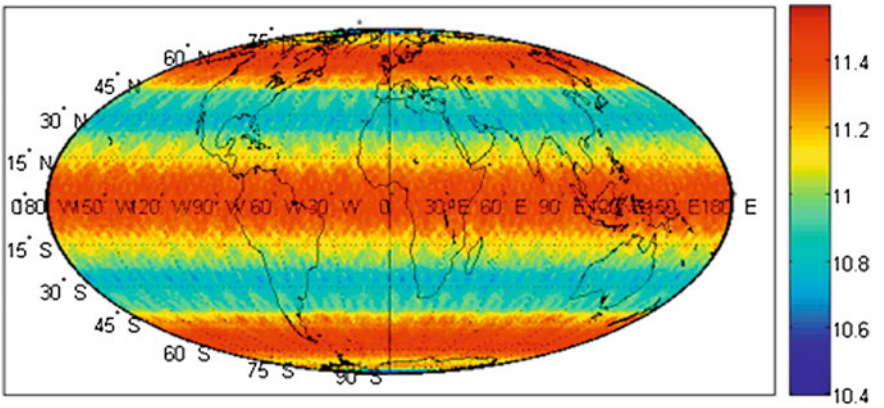


Fig. 4.10 Positioning precision simulation for sphere with orbit of 30,000 km

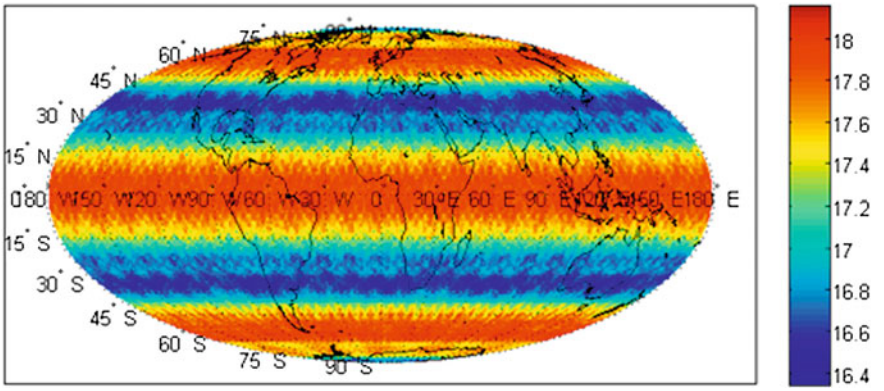


Fig. 4.11 Positioning precision simulation for geostationary orbit sphere

2. The average positioning error is less than 6.6 m in the sphere of 24,000 km, and the average positioning error is less than 11.6 m in the sphere of 30,000 km. The average positioning error is less than 18.2 m in geostationary orbit sphere.
3. The average positioning error can achieve more than 18.2 m between the area of MEO orbit sphere and geostationary orbit sphere, if the transmitting beam width of $\pm$ 60° is adopted.

4.7 Conclusions

Based on the MEO constellation and antenna direction, this paper proposes a navigation signal for space applications. Theoretically analysis and simulation shows that:

1) The antenna half angle is about ±60° in order to receive the navigation signal outside the shadow of the earth and meet the requirement of visibility and GDOP.
2) The proposed antenna gain is designed to reach its maximum when the transmitting half angle is 50°. The received signal power is −165.7, −166.6 and −167.3 dBW for the spacecraft in the sphere with radius of 24,000, 30,00 and 35,786 km respectively within the visible area in which the transmitting half angle is 13.2° ∼ 60°.
3) When the proposed antenna width and pattern is used, the UERE is 1.8, 1.87 and 1.93 m for the spacecraft in the sphere with radius of 24,000, 3,000 and 35,786 km respectively within the visible area in which the transmitting half angle is 13.2° ∼ 60°.
4) Based on the compass constellation, the proposed spacecraft navigation signal has more than 11 available satellites and less than 17.3 m positioning error in the geostationary orbit sphere.
5) The available satellite number can achieve more than 11 between the area of MEO orbit sphere and geostationary orbit sphere.
6) The average positioning error can achieve more than 18.2 m between the area of MEO orbit sphere and geostationary orbit sphere.

This paper studied the spacecraft navigation, and proposed a width beam spacecraft navigation signal, which meet the navigational requirement of constellation and antenna direct. Furthermore, the implementation is accepted by the existing MEO constellation, such as the antenna design and transmitting power. The proposed signal provides a possible solution for the space navigation signal design in the Chinese compass construction.

References

1. Birmingham WP, Miller BL, Stein WL (1983) Experimental results of using the GPS for landsat-4 onboard navigation. Navig: J Inst Navig 30(3):244–251
2. Kronman JD, McElroy T (1994) Considerations for the application of GPS in satellites: GPS system trade study. In: Proceedings of ION GPS-94
3. Wu S (1996) MicroGPS for orbit determination of earth satellites. In: Proceedings of ION NTM 1996, pp 509–516
4. Cohen CE (1994) Space flight tests of attitude determination using GPS. Int J Satell Commun 12:427–433
5. Kronman JD (1996) Experience using GPS for orbit determination of geosynchronous satellite. In: Proceedings of ION GPS 2000, pp 1622–1626
6. Michael C (2002) Results from the GPS flight experiment on the high earth orbit AMST OSCAR-40 spacecraft. In: Proceedings of ION GPS 2002, pp 122–133
7. Global Positioning System Standard Positioning Service Performance Standard (2001)
8. Honglei Q, Minmin L (2008) Research on positioning of high earth orbital satellite using GNSS. Chin J Space Sci 28(4):316–325

9. He Q, Sun Q (2010) A new solution of autonomous navigation for GEO satellites based on GNSS. J Spacecraft TT & C Technol 29(1):7–11
10. Betz JW, Kolodziejski KR (2009) Generalized theory of code tracking with an early-late discriminator part II: noncoherent processing and numerical results. IEEE Trans Aerosp Electron Syst 45(4):1551–1564

Chapter 5
Research and Analysis of BOC(1,1) Multipath Mitigation Performance Based on ASPeCT

Hang Ruan, Lei Zhang, Feng Liu and Zhou Zheng

Abstract BOC(1,1) has a high accuracy of code tracking and perfect multipath mitigation performance as the main modulation of the next generation of GPS and the European GALILEO. Autocorrelation Side—Peak Cancellation Technique (ASPeCT) is specially design for BOC(1,1) code tracking loop and further improves the BOC(1,1) code tracking accuracy and multipath mitigation performance. In this paper, the multipath mitigation performance of ASPeCT applied to BOC(1,1) is researched and analyzed in case of the infinite bandwidth and limited bandwidth in consideration of the real channel properties. In the end, it gains the optimal correlating space and its corresponding minimum receiver bandwidth.

Keywords ASPeCT · Multipath mitigation · BOC(1,1) · Bandwidth limited

5.1 Introduction

The newly designed modulation BOC(1,1) [1] will be applied to the modernization GPS and Galileo system. On one hand, the spectral of BOC(1,1) splits from the central frequency point which can propagate together with the origin BPSK signal at the same frequency point. On the other hand, the new signal will improve the pseudo random code tracking accuracy and the performance of the multipath mitigation. Especially the improvement of the multipath mitigation BOC(1,1) brings is particularly significant because of the sharper main peak of the autocorrelation compared with the traditional BPSK signal. However, the multi-peak characteristic may make the loop lock on the side peak and bring a huge tracking error.

H. Ruan (✉) · L. Zhang · F. Liu · Z. Zheng
School of Information and Electronics, Beijing Institude of Technology, Beijing, People's Republic of China
e-mail: ruanhang_bit@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_5,

At present, the primary BOC(1,1) signal processing approaches include Bump-Jump [2], the BPSK-LIKE method [3], dual estimation tracking (DET) [4], the ASPeCT [5] which will be analyzed in this paper and so on. All of these approaches are able to solve the problem locking on the side peak. But the autocorrelation of BPSK-LIKE method is more or less the same as the autocorrelation of BPSK signal and it loses the outstanding multipath mitigation of BOC(1,1). The DET method needs two loop to track the signal so that it is complex to realized. The Bump-Jump method tracks the BOC(1,1) signal directly but how to detect whether the loop lock on the side peak is difficult. The ASPeCT further improve the performance because of the autocorrelation is changed and the main peak is much sharper. However, the detail analysis especially the limited bandwidth multipath performance has not made. So in this paper, it analyzes the multipath mitigation performance of ASPeCT and makes a comparison between ASPeCT and BOC(1,1) based on multipath error envelop [6]. Especially, the limited bandwidth performance is analyzed and get out the optimum correlation space and its corresponding minimum receiver bandwidth.

The paper is organized as follow: in the first section, the ASPeCT is introduced. In section two, the infinite bandwidth multipath mitigation performance of ASPeCT is analyzed. In section three, the limited bandwidth performance of ASPeCT is analyzed. In section four, simulation and comparison are done for ASPeCT and BOC(1,1) and gain the conclusion.

5.2 ASPeCT Scheme

BOC(1,1) is widely used as civilian navigation signals in the next generation of GNSS. It is the major component of the GALILEO system E1 frequency point and modernization GPS L1C frequency point. The pilot channel signal expression is as Eq. (5.1) where P is the signal power, c(t) is pseudo random code, f_p is the pilot signal carrier frequency, f_d is the Doppler frequency, φ_0 is the initial carrier phase, n(t) is additive Gaussian white noise. sc(t) is the square subcarrier, including two form of sine-phase and cosine-phase. In this paper, only the sine-phase BOC(1,1) is researched. The sine phase square subcarrier expression is as Eq. (5.2).

$$s(t) = \sqrt{2P}c(t)sc(t)\cos\left(2\pi\left(f_p + f_d\right)t + \varphi_0\right) + n(t) \tag{5.1}$$

$$sc(t) = \mathrm{sign}(\sin(f_{sc}t)) \tag{5.2}$$

where f_{sc} is the square subcarrier frequency. Compared with the traditional BPSK modulation, the newly approach modules an extra square subcarrier together with the pseudo random code. So the autocorrelation function is no longer a simple triangle function and its expression is as Eq (5.3) In the approach of ASPeCT, the cross correlation between BOC(1,1) and PRN code is also used to build the composite correlation and its expression is as Eq. (5.4) where τ is the code

$$R_{BOC}(\tau) = \text{tri}(\tau) - \frac{1}{2}\left[\text{tri}\left(\tau + \frac{1}{2}\right) + \text{tri}\left(\tau - \frac{1}{2}\right)\right] \tag{5.3}$$

$$R_{BOC/PRN} = \frac{1}{2}\left[\text{tri}\left(\tau + \frac{1}{2}\right) - \text{tri}(\tau - \frac{1}{2})\right] \tag{5.4}$$

delay and tri (τ) is the standard triangle function whose width and height are both with value of 1. The autocorrelation function of BOC(1,1) and the cross-correlation function between BOC(1,1) and PRN code are illustrated in Fig. 5.1.

We can get the conclusion from the figure that the autocorrelation's main peak of BOC(1,1) is sharper than that of the BPSK signal so that the multipath mitigation performance will be improved. However, there are another two side peaks at the delay of 0.5 chip. There are two peaks in the cross correlation function be tween BOC(1,1) and the PRN at the delay of 0.5 chip. The magnitude of the peak is the same as the magnitude of the autocorrelation's side peak.

Autocorrelation Side—Peak Cancellation Technique(ASPeCT) is a designed approach on BOC(1,1) signal code tracking. It makes use of the situation that the location and amplitude of the BOC(1,1) autocorrelation's side peaks and the location and amplitude of the cross correlation are the same. By subtracting the cross correlation from the autocorrelation of BOC(1,1), the side peak of the autocorrelation can be diminished and the main peak is sharper at the same time which bring the performance improvement.

There are two approaches to subtract the cross correlation function from the autocorrelation function which are modulus value subtraction and power

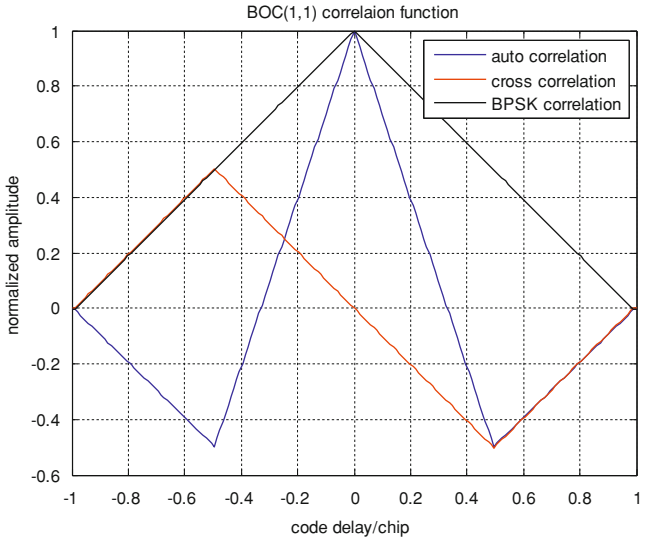


Fig. 5.1 BOC(1,1) correlation function The *blue*, *red* and *black curves* are the ACF of BOC(1,1), the cross correlation function of BOC(1,1) and the ACF of BPSK respectively

subtraction. These two methods are corresponding to the power discriminator and modulus discriminator of EMLP respectively. The composite correlation functions are as Eqs. (5.5) and (5.6)

$$R^2_{ASPeCT}(\tau) = R^2_{BOC}(\tau) - \boldsymbol{\beta} R^2_{BOC.PRN}(\tau) \tag{5.5}$$

$$R_{ASPeCT}(\tau) = |R_{BOC}(\tau)| - \beta \left| R_{BOC/PRN}(\tau) \right| \tag{5.6}$$

where β is the adjustment coefficient which is used to adjust the shape of the correlation function. If we set $\beta = 1$, the two composite correlation functions are illustrated in Fig. (5.2). We can get the conclusion that there is a positive peak and two negative peaks in the correlation of ASPeCT. As the discriminator will not get the modulus value of the correlator output so that the side peak will not be tracked. In the traditional EMLP discriminator, the early correlator and the late correlator are set near the main peak of the autocorrelation. The power discriminator and the modulus discriminator are as Eqs (5.7) and (5.8)

$$D_p(\varepsilon) = \frac{\left[\left(I_E^2 + Q_E^2\right)_{B/B} - \left(I_E^2 + Q_E^2\right)_{B/P} \right] - \left[\left(I_L^2 + Q_L^2\right)_{B/B} - \left(I_L^2 + Q_L^2\right)_{B/P} \right]}{\left(I_E^2 + Q_E^2\right)_{B/B} + \left(I_L^2 + Q_L^2\right)_{B/B}} \tag{5.7}$$

$$D_{AM}(\varepsilon) = \frac{\left[\left(I_E^2 + Q_E^2\right)^{1/2}_{B/B} - \left(I_E^2 + Q_E^2\right)^{1/2}_{B/P} \right] - \left[\left(I_L^2 + Q_L^2\right)^{1/2}_{B/B} - \left(I_L^2 + Q_L^2\right)^{1/2}_{B/P} \right]}{\left(I_E^2 + Q_E^2\right)^{1/2}_{B/B} + \left(I_L^2 + Q_L^2\right)^{1/2}_{B/B}} \tag{5.8}$$

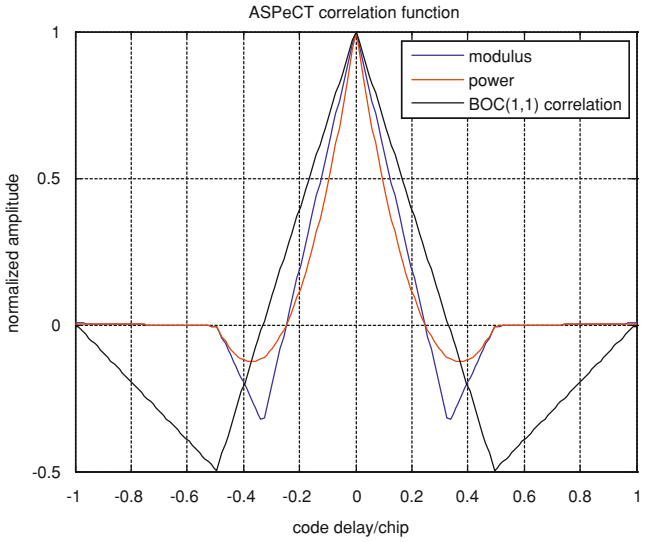


Fig. 5.2 ASPeCT correlation function. The *blue*, *red* and *black curves* are the module, power ASPeCT correlation function and BOC(1,1) ACF

As the cross correlation output makes no contribution to the discriminator output in power, only the component of the autocorrelation output is used to normalize. So we can get the conclusion from the figure that the multipath mitigation performance will be outstanding compared with the traditional method when the correlation space is same. The multipath performance will be analyzed in the next section based on modulus value discriminator in consideration of the practical application.

5.3 Analysis Based on Multipath Error Envelop

If there are *m* multipath signals received together with the direct BOC(1,1) signal by GNSS receiver without consideration of the limited bandwidth of the receiver and the channel, the base band receiving signal can be written as Eq. (5.9) where c(t) is the PRN, sc(t) is the square subcarrier of BOC(1,1), P, τ_0, φ are the direct signal power, the direct signal delay and the direct signal carrier phase respectively. α_i, $\Delta\tau_i$, $\Delta\varphi_i$ is the power, delay and carrier phase of the ith multipath signal corresponding to the direct signal. In this paper, only 1 multipath is put into the received signal when analyzing the performance of ASPeCT. So the ASPeCT composite correlation output between the received signal and local replica can be written as Eq. (5.10) if only one multipath exists.

$$\begin{aligned} r(t) = {} & \sqrt{2P}c(t-\tau_0)sc(t-\tau_0)\cos(\varphi) \\ & + \sqrt{2P}\sum_{i=1}^{m}\alpha_i c(t-\tau_0-\Delta\tau_i)sc(t-\tau_0-\Delta\tau_i)\cos(\varphi+\Delta\varphi_i) \end{aligned} \tag{5.9}$$

$$\begin{aligned} R_{rs}(\tau) = {} & \sqrt{2P}TR(\tau-\tau_0)\cos(\varphi_0-\varphi_c) \\ & + \sqrt{2P}\alpha_1 TR(t-\tau_0-\Delta\tau_1)\cos(\varphi_0-\Delta\varphi_i-\varphi_c) \end{aligned} \tag{5.10}$$

The local replica completely synchronizes with the received signal if it is $\tau = \tau_0$. T is the integration accumulation length. $R(\tau)$ is the composite correlation function of ASPeCT which can be written as Eq. (5.11) where T_c is the pseudo random code period, $R_{ASPeCT}(\tau)$ is the even symmetry function with respect to $\tau = 0$. The composite correlation function of ASPeCT is as Fig. 5.3. If one multipath signal is received at the same time direct signal is received. In the figure, there are two situations of the carrier phase which are $\varphi_0 - \Delta\varphi_1 - \varphi_c = \pi$ and $\varphi_0 - \Delta\varphi_1 - \varphi_c = 0$ respectively. The amplitude of the multipath signal is the half of the direct signal. The multipath delay is 0.1 chips

$$R_{ASPeCT}(\tau) = \begin{cases} -\frac{2}{T_c}\left(\tau+\frac{1}{2}T_c\right), -\frac{T_c}{2} \le \tau \le -\frac{T_c}{3}; \frac{4}{T_c}\left(\tau+\frac{1}{4}T_c\right), -\frac{T_c}{3} \le \tau \le 0 \\ -\frac{4}{T_c}\left(\tau-\frac{1}{4}T_c\right), 0 \le \tau \le \frac{T_c}{3}; \frac{2}{T_c}\left(\tau-\frac{1}{2}T_c\right), \frac{T_c}{3} \le \tau \le \frac{T_c}{2} \end{cases} \tag{5.11}$$

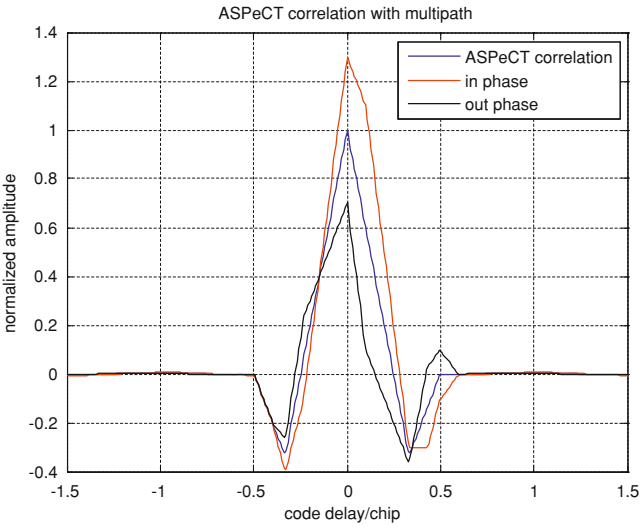


Fig. 5.3 ASPeCT correlation function. The *blue*, *red* and *black curves* are ASPeCT correlation function, and that with in phase multipath and out phase multipath

The discriminator of the code tracking loop based on ASPeCT is the same as the traditional EMLP discriminator. The early, prompt and the late correlators are set in turn with the space of d/2 (d < 1/3chip). The EMLP discriminator output is as Eq. (5.12) where $R_{ASPeCT}(\tau)$ is the normalized correlation function of ASPeCT

$$D_{AM}(\varepsilon) = R_{ASPeCT}\left(\varepsilon - \frac{d}{2}\right) - R_{ASPeCT}\left(\varepsilon + \frac{d}{2}\right) \tag{5.12}$$

When multipath signal exists, the estimating error of the real direct signal code phase is as Eq. (5.13). We can get the result from the expression that the multipath error will be the maximum if the carrier phase error is 0 or π corresponding to one multipath. So the multipath error envelop is composed of the two curve of the limiting case. In this paper, the multipath delay analyzed is from 0 to 1.2 chips for the reason that the error is 0 when the delay extends 1.2 chips. The multipath error envelop of ASPeCT and BOC(1,1) are both illustrated in Fig. 5.4

$$ME(\Delta\tau) = \begin{cases} \alpha_1 \Delta\tau \cos\Delta\varphi_i, 0 < \Delta\tau \le \frac{d}{2}; \frac{d}{2}\alpha_1 \cos\Delta\varphi_i, \frac{d}{2} < \Delta\tau \le \frac{T_c}{3} - \frac{d}{2} \\ -\frac{1}{4}\alpha_1\left(\Delta\tau - \frac{T_c}{3} - \frac{d}{6}\right)\cos\Delta\varphi_i, \frac{T_c}{3} - \frac{d}{2} < \Delta\tau \le \frac{T_c}{3} + \frac{d}{2}; \\ -\frac{1}{4}\alpha_1 \cos\Delta\varphi_i, \frac{T_c}{3} + \frac{d}{2} < \Delta\tau \le \frac{T_c}{2} - \frac{d}{2}; \\ \frac{1}{2}\alpha_1\left(\Delta\tau - \frac{T_c}{2} - \frac{d}{2}\right)\cos\Delta\varphi_i, \frac{T_c}{2} - \frac{d}{2} < \Delta\tau \le \frac{T_c}{2} + \frac{d}{2}; \end{cases} \tag{5.13}$$

In this paper, the ratio between multipath amplitude and direct signal multipath (MDR) is defined as Eq. (5.14) and the MDR is 6 dB.

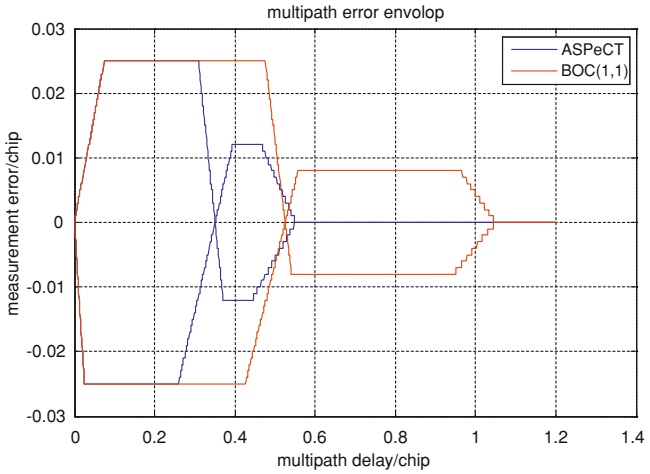


Fig. 5.4 Infinite bandwidth multipath error envelop. The *blue curve* and the *red curve* are multipath error envelop of ASPeCT and BOC(1,1) respectively

$$\mathrm{MDR} = 10\log\left(\mathrm{AM}_{\mathrm{LOS}}/\mathrm{AM}_{\mathrm{mul}}\right) \tag{5.14}$$

Multipath error envelop reflects the tracking performance when the multipath error exists, it is right to evaluate the multipath mitigation performance according to the multipath error envelop. Essentially, the multipath error envelop is composed of the in-phase and inverting phase multipath error. In the middle of the curve, it is the multipath error with any other phase. By comparing multipath error envelop of the BOC(1,1) with narrow correlation and the ASPeCT with EMLP, we can gain the conclusion as follow:

1. The multipath mitigation performance of ASPeCT is the same as the BOC(1,1) narrow correlation when the delay of multipath is less 0.2 chip
2. The multipath mitigation performance of ASPeCT is greatly improved when the delay is longer the 0.2 chip.

In total, the ASPeCT has a better performance for middle and long delay multipath.

5.4 Bandwidth Limited Simulation and Analysis

In fact, the navigation signal the satellite sends is bandwidth limited. The signal channel and the receiver frond-end are also bandwidth limited. So the received signal is bandwidth limited which can be expressed as an infinite bandwidth signal propagating through a band pass filter. As the BOC subcarrier move the origin spectrum apart from the centre frequency point, the bandwidth of the signal is

wider than BPSK signal, the bandwidth analyzed in the paper is $B = 2f_s + 2f_c$ and $B = 2f_s + 10f_c$ respectively where f_s is subcarrier frequency and f_c is the pseudo random code frequency. The following Fig. 5.5 illustrates the correlation of ASPeCT of the two bandwidths.

According to the two figures, we can gain the conclusion as follows

When the bandwidth B fulfills $B = 2f_s + 10f_c$, as most part of the signal spectrum is involved in the receiving band, the bandwidth limited correlation function and the infinite bandwidth correlation function is more or less the same.

When the bandwidth B fulfills $B = 2f_s + 2f_c$, as many harmonic components are diminished by the bandwidth limited filter, the main peak of the correlation function become smooth and wide. The performance of multipath mitigation is degenerated.

In the simulation, there is a multipath together with the direct signal received by the receiver. The front-end bandwidth of the receiver analyzed in the paper is $B = 2f_s + 10f_c$ and $B = 2f_s + 2f_c$ respectively. Compared with the situation of infinite bandwidth, the multipath error envelop of ASPeCT and BOC(1,1) with narrow correlation is as Fig. 5.6 respectively.

According to the two figures conclusion is gained as follow

When the bandwidth B fulfills $B = 2f_s + 10f_c$, both of ASPeCT and BOC(1,1) narrow correlation have nearly the same performance as that in condition of infinite bandwidth. Furthermore, for short delay the multipath mitigation is more or less the same, for long delay ASPeCT is better.

When the bandwidth B fulfills $B = 2f_s + 2f_c$, the bandwidth is as the same as code frequency. The multipath mitigation performance of both ASPeCT and BOC(1,1) is deteriorated. The distortion of BOC(1,1) is more serious and ASPeCT has better performance than BOC(1,1) for both long delay and short delay.

The effective multipath range is also researched in order to further analyze the influence of the limited bandwidth on ASPeCT. The comparison is also done

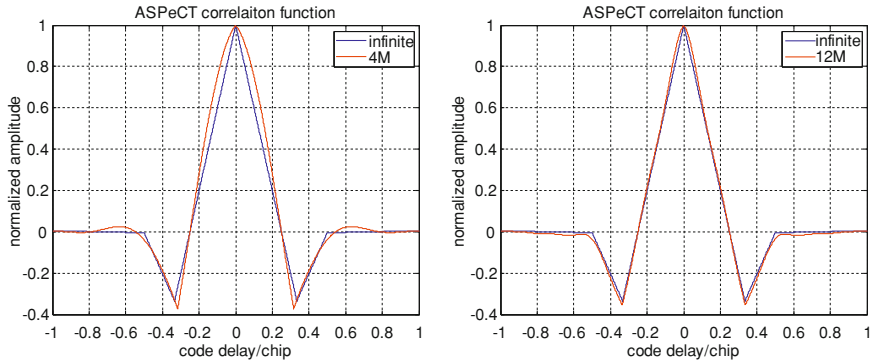


Fig. 5.5 ASPeCT correlation function with different bandwidth. The *blue curve* and the *red curve* of the *left* figure are the correlation function of infinite bandwidth and 4 MHz. The *blue curve* and the *red curve* of the *right* figure are the correlation function of infinite bandwidth and 12 MHz

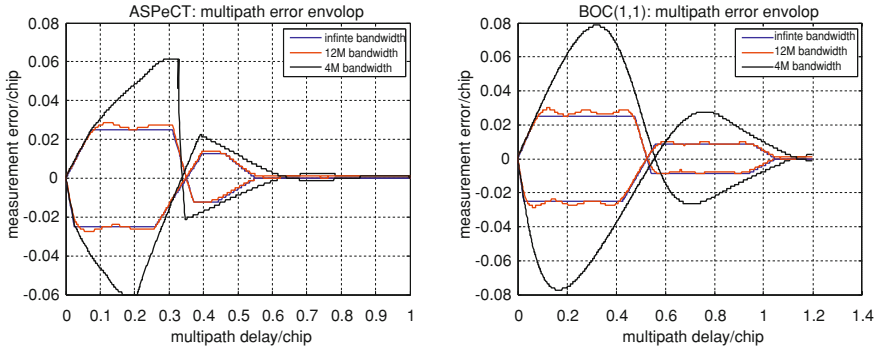


Fig. 5.6 Multipath error envelop. The *blue*, *red* and *black curves* of the *left* figure are ASPeCT multipath error envelop with the bandwidth of infinite, 12 and 4 MHz. The *blue*, *red* and *black curves* of the left figure are BOC(1,1) multipath error envelop with the bandwidth of infinite, 12 and 4 MHz

between BOC(1,1) narrow correlation and ASPeCT. When only one multipath signal is received with the direct signal and the correlation space d is 0.1 chip, the effective multipath range of ASPeCT and BOC(1,1) with respect to different bandwidth is illustrated in Fig. 5.7 respectively.

According to the figures, conclusion is gained as follows:

1. When the front-end bandwidth is narrow, the average error of both methods is maximum at the delay of 0.4 chip. ASPeCT has the performance much better than BOC(1,1) as the bandwidth increases.
2. When the front-end bandwidth is relatively wide, the performance is more or less the same as the infinite bandwidth. As the same time, ASPeCT is still better than BOC(1,1) autocorrelation but the priority is not that obvious as that with narrow bandwidth.

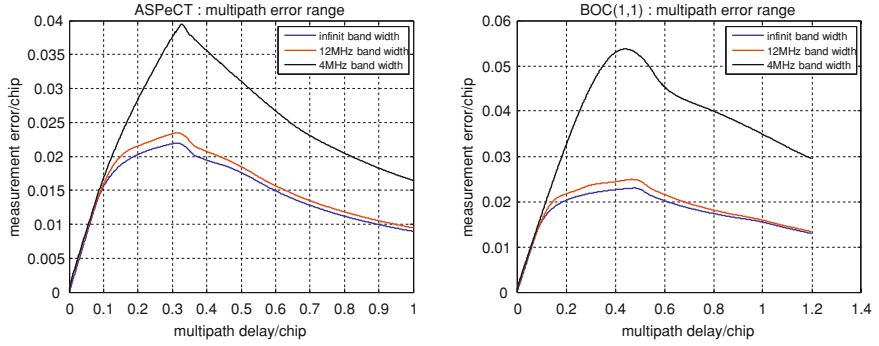


Fig. 5.7 Multipath error range. The *blue*, *red* and *black* curves of the *left* figure are ASPeCT multipath error range with the bandwidth of infinite, 12 and 4 MHz. The *blue*, *red* and *black curves* of the *left* figure are BOC(1,1) multipath error range with the bandwidth of infinite, 12 and 4 MHz

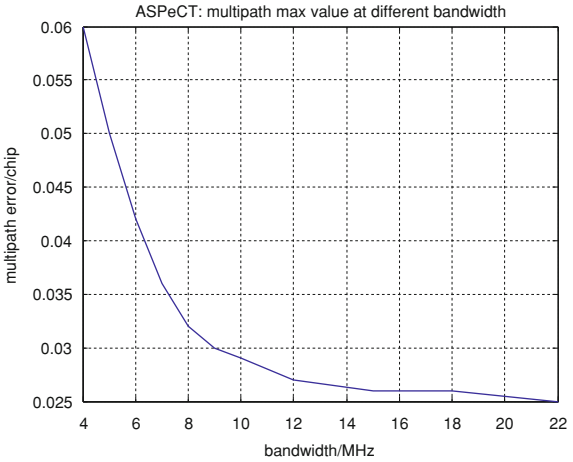


Fig. 5.8 ASPeCT multipath error different bandwidth

We can also gain the conclusion that if the bandwidth is as wide as 1/d and the correlation space is d, the multipath mitigation performance is similar to that of infinite bandwidth. There is no effect to increase the bandwidth continuously. The multipath error extremism of different bandwidth with 0.1 chip correlation space is as Fig. 5.8.

As the signal the satellite sends is bandwidth limited, no matter how wide the receiver bandwidth is, the performance cannot be better. Indeed, the signal sent is more or less the 20 MHz wide. The multipath error of different correlation space corresponding to 20 MHz bandwidth is simulated and the result is as Fig. 5.9.

According to the simulation, it is clearly that the performance of multipath mitigation will be the best if the correlation space is 0.05 chips to 0.1 chips. The

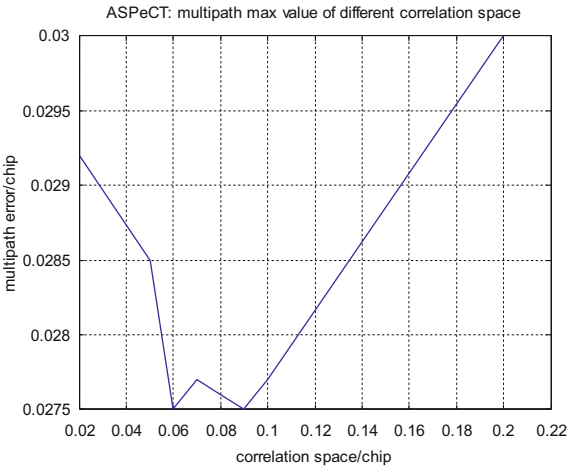


Fig. 5.9 ASPeCT multipath error different bandwidth

correlation space can never be smaller because of noise and bandwidth. In consideration of the bandwidth limited condition and the noise influence, the correlation space is usually chosen to be 0.1 chips.

5.5 Conclusion

The paper mainly analyzes the multipath mitigation performance of ASPeCT which is design for BOC(1,1) and make a comparison with the method of BOC(1, 1) narrow correlation. The situation of infinite bandwidth and limited bandwidth are both analyzed. The bandwidth limited autocorrelation function is gained from frequency domain. Simulation is done to get the bandwidth limited performance of ASPeCT and BOC(1,1) narrow correlation and the optimum correlation space and its corresponding bandwidth is gained from the simulation.

References

1. Betz JW (1999) Binary offset carrier modulations for radio navigation. In: Proceedings of ION Technical meeting, pp 639–648
2. Fine P, Wilson W (1999) Tracking algorithm for GPS offset carrier signals. In Proceedings of U.S institute of navigation NTM conference, San Diego, CA, pp 671–676
3. Martin N, Leblond V, Guillotel G, Heiries V (2003) BOC(x, y) signal acquisition techniques and performances. In Proceedings of ION GPS/GNSS 2003, 9–12 September 2003, Portland, pp 188–198
4. Martin N, Guichon H (2009) "BOC ambiguity resolution aided by virtual sub-carrier tracking for multi-path mitigation" GNSS Signals 2009, December 2009
5. Julien O, Macabiau C, Cannon ME et al (2007) ASPeCT: unambiguous sine-BOC (n, n) acquisition/tracking technique for navigation applications aerospace and electronic systems 43(1):150–162
6. Irsigler M (2005) Criteria for GNSS multipath performance assessment [C] In: Proceedings of the 18th international technical meeting of the satellite division of the institute of navigation, pp 166–177

Chapter 6
Study on MSK Modulation for S-band

Fengjiao Wang, Dazhi Zeng and Rui Li

Abstract With the development of global navigation satellite system (GNSS), the number of signals transmitted in L-band has been increasing. Exploiting new frequency bands for radio navigation satellite service (RNSS) is quite necessary. S-band has received widely attention since the performance of positioning and timing will be benefited significantly by joint positioning of S-band and L-band. The minimum shift keying—binary offset carrier (MSK-BOC) modulation was proposed to be used for Beidou-2 S-band here. The compatibility between Beidou-2 and Beidou-1/Globalstar system was firstly evaluated, and then signal's code tracking error and multipath error envelope were analyzed. Results indicate that the performance in terms of compatibility and tracking precision of MSK-BOC modulation in S-band is good. The research will provide an instructive reference for the signal design and engineering implementation of Beidou-2.

Keywords S-band · MSK modulation · Compatibility · Interference · Code tracking precision

6.1 Introduction

Satellite navigation system is a country's important information infrastructure and strategic facilities, which plays an important part in the fields of military, economic and social life. With the construction and development of global navigation satellite system (GNSS), the number of signals transmitted in L-band has been increasing significantly. At the same time, by joint positioning of different bands, the performance of GNSS can be improved, so it is quite necessary to exploit new

F. Wang · D. Zeng (✉) · R. Li
School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
e-mail: dazhi@bit.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_6,

bands used for radio navigation satellite service (RNSS). China is vigorously developing its independent Beidou-2 satellite navigation system. The status quo is that signals in L-band have been identified preliminarily while for new navigation bands, the signal design work has just begun. Compared with other bands, S-band (2483.5–2500 MHz) can be utilized to combine mobile communication with navigation positioning service by a universal antenna to reduce the terminal's hardware complexity [1]. The receiver's ionosphere correction ability and the precision of timing and positioning can be improved by Meta-GNSS of L-band and S-band. The precision and robustness of precise point positioning can also be improved, because the joint positioning of S-band and L-band is conductive to solve the integer ambiguity. Therefore, using S-band for satellite radio navigation has drawn attention of a growing number of countries and regions. S-band has already been used for the downlink of China's Beidou-1 navigation system. The areal satellite positioning and short message communication have been realized by Beidou-1 which means China's research of radio determination satellite service (RDSS) in S-band is in the lead. In order to improve S-band's regional advantage and reserve resources for Beidou-2 system, research on Beidou-2 signal structure design especially the signal's modulation mode in S-band is quite important.

The minimum shift keying (MSK) modulation is a kind of constant envelope modulation with continuous phase varying, which will not bring nonlinear distortion when through the high power amplifier. Compared with binary phase shift keying (BPSK) modulation, MSK can change the pulse shape of spreading codes. The influenced power spectrum density (PSD) has higher power in the main lobe which makes MSK to get better performance in certain bandwidth. While the lower power in side lobe make less influence on other signals in the same band [3]. Furthermore, MSK can be realized and combined with the binary offset carrier (BOC) modulation simply. Many communication and navigation systems are already in S-band (2,483.5–2,500 MHz), therefore a new kind of modulation with good performance is needed for RNSS in S-band to realize high compatibility. Study on MSK modulation in S-band will provide an instructive reference for Beidou-2 S-band's design and engineering implementation.

6.2 Mathematical Model of MSK Modulation

6.2.1 Time Domain Model

The MSK modulation signal in time domain can be expressed as follows [2]:

$$s_{MSK}(t) = a_k(t)\cos(\pi t/T_c)\cos 2\pi f_L t + b_l(t)\sin(\pi t/T_c)\sin 2\pi f_L t$$

where a_k and b_l are the spreading code sequences of in-phase and quadrature branch separately, and compared with a_k, b_l delays by half a code chip period; T_c is the code chip duration; f_L is the carrier central frequency. Compared with the

QPSK modulation, MSK changes the shape of spreading code by multiplying code sequence with cosine signal.

When the data and pilot signals are all modulated by rectangle spreading codes, without considering the navigation data, converting the code sequence's pulse shaping to single code's pulse shaping, we can get the baseband MSK-R signal as:

$$s_{MSK-R}(t) = \sum_{k=-\infty}^{\infty} a_k \cos[\pi(t - kT_c)/T_c] \cdot rect[(t - kT_c)/T_c] + j \sum_{l=-\infty}^{\infty} b_l \sin[\pi(t - T_c/2 - lT_c)/T_c] \cdot rect[(t - T_c/2 - lT_c)/T_c]$$

where

$$rect(x) = \begin{cases} 1, & x \in [0, 1] \\ 0, & otherwise \end{cases}$$

If the spreading codes have already been modulated by square sub-carrier, then the result of MSK modulation is to do cosine half cycle shaping to the pulse shape of square sub-carrier. We can get the baseband MSK-BOC signal as:

$$s_{MSK-BOC}(t) = \sum_{k=-\infty}^{\infty} a_k rect[(t - kT_c)/T_c] \cdot \cos[2\pi(t - kT_s/2)/T_s] \cdot sign[\sin(2\pi f_s(t - kT_c))] + j \sum_{l=-\infty}^{\infty} b_l rect[(t - T_s/4 - lT_c)/T_c] \cdot \sin[2\pi(t - T_s/4 - lT_s/2)/T_s] \cdot sign[\sin(2\pi f_s(t - T_s/4 - lT_c))]$$

where T_s is the period of square sub-carrier; f_s is the frequency of square sub-carrier; $sign(\bullet)$is the sign function.

6.2.2 Power Spectral Density Model

By taking the discrete Fourier transform (DFT) of the MSK modulation signal in the time domain, the PSD of MSK-R can be expressed as [3]:

$$G_{MSK-R}(f) = \frac{8f_c^3 \cos^2(\pi f/f_c)}{\pi^2 \left(f_c^2 - 4f^2\right)^2}$$

where f_c is the chip rate.

Similarly, the PSD of MSK-BOC can be written as:

$$G_{MSK-BOC}(f) = \begin{cases} \frac{2f_s^2 f_c \sin^2(\pi f/f_c)}{\pi^2 (f_s^2 - f^2)^2}, & k \text{ is even} \\ \frac{2f_s^2 f_c \cos^2(\pi f/f_c)}{\pi^2 (f_s^2 - f^2)^2}, & k \text{ is odd} \end{cases}$$

where $k = 2f_s/f_c$ is the index of MSK-BOC modulation.

6.3 Compatibility of S-band

Although developing navigation services of Beidou-2 navigation system in S-band has great advantages, the interference is unavoidable. Here, several kinds of modulation mode such as MSK(4), BPSK(4), BPSK(8), BOC(4,4), MSK-BOC(4,4) and MSK-BOC(4,2) are put forward to be used in S-band of Beidou-2 navigation system. Through calculating the spectral separation coefficient (SSC) [6], code tracking sensitivity coefficient ($\eta_{j,s}$) [7] and effective carrier-to-noise loss (C/N_0) caused by interference, the acceptance of interference between Beidou-2 and S-band's other systems such as Beidou-1 and Globalstar system can be estimated. In other words, whether the compatibility in S-band is good can be estimated.

6.3.1 Interference with Beidou-1 System

Beidou-1 system uses S-band for its downlink communications between the satellite and user terminals. The central frequency of Beidou-1 signal in S-band is 2491.75 MHz and the modulation mode is BPSK(4). In order to analyze the compatibility, suppose that:

- The central frequency of Beidou-2 system in S-band is 2491.028 MHz.
- There are at most 12 Beidou-2 satellites in the space at the same time while for Beidou-1 there are only 2 and the interference of every satellite is the same.
- The power flux density of every Beidou-1 or Beidou-2 satellite is the threshold value of -126 dBW/m^2/MHz, which represents the greatest interference.
- The antenna gain is 3 dB and the effective area is -29.4dBm^2.
- The noise PSD is –203.83 dBW/MHz, while the intra-system interference is always below -222 dBW/MHz [4].
- For the supposed modulation, the receiver band is the corresponding signal width of main lobe. If the main lobe width is larger than 16.368 MHz, choose 16.368 MHz as the receiver band.

The simulation results are shown in Tables 6.1 and 6.2. We can find that of all the modulations, the interference of MSK-BOC(4,2), MSK-BOC(4,4) and BOC(4,4) is very small.

6.3.2 Interference with Globalstar System

Globalstar mainly provides communication services for area between north latitude 70° and south latitude 70°, every service area is covered by 2–4 satellites. If

Table 6.1 Interference from beidou-1 to beidou-2

Modulation mode	Receiver bandwidth MHz	*SSC* dB/Hz	$\eta_{j,s}$ dB/Hz	I_0 dBW/Hz	C/N_0 dB
MSK(4)	12.276	−68.44	−68.29	−206.93	1.71
BPSK(4)	8.184	−67.96	−67.49	−208.21	1.33
BPSK(8)	16.368	−70.25	−69.50	−207.49	1.53
BOC(4,4)	16.368	−72.43	−73.15	−209.67	0.99
MSK-BOC(4,4)	16.368	−72.89	−74.16	−210.13	0.90
MSK-BOC(4,2)	12.276	−76.41	−77.02	−214.90	0.32

Table 6.2 Interference from beidou-2 to beidou-1

Modulation mode	Receiver bandwidth MHz	*SSC* dB/Hz	$\eta_{j,s}$ dB/Hz	I_0 dBW/Hz	C/N_0 dB
MSK(4)	8.184	−68.32	−67.87	−200.79	4.75
BPSK(4)	8.184	−67.96	−67.49	−200.43	4.99
BPSK(8)	8.184	−69.98	−69.53	−202.45	3.72
BOC(4,4)	8.184	−73.86	−73.50	−206.33	1.91
MSK-BOC(4,4)	8.184	−74.39	−74.01	−206.86	1.73
MSK-BOC(4,2)	8.184	−76.90	−77.00	−209.37	1.06

S-band is used in Beidou-2 system, combining Beidou-2 and mobile communication will have a lot of advantages, but the interference is also unavoidable.

Globalstar uses the 2483.5–2500 MHz for the downlink communications between satellite and user terminals with 2491.75 MHz as the central frequency. The system uses multi-beam antennas to allow frequency multiplexing. In every beam, the 16.5 MHz bandwidth is divided into 13 FDM channels, each is 1.23 MHz wide. CDMA with a chipping rate of 1.2288Mcps is implemented inside every FDM channel. Before modulated of the carrier, the Globalstar signal is filtered by a Nyquist-square-root raised-cosine filter with roll-off factor $\rho = 0.2$. The use of the raised cosine filter yields the following power spectral density for a given *kth* FDMA channel in baseband [5]:

$$G_{GLOB}(f) = \left(\sum_{k=-6}^{6} G_{SRC}^{k}(f) \right) / 13$$

where

$$G_{SRC}^{k}(f) = \begin{cases} 1/f_c & if\ |f - kB| \le f_c(1-\rho)/2 \\ 0 & if\ |f - kB| \le f_c(1+\rho)/2 \\ g(f) & if\ f_c(1-\rho)/2 \le |f - kB| \le f_c(1+\rho)/2 \end{cases}$$

$$g(f) = \frac{1}{2f_c}\left(1 + \cos\left(\frac{\pi}{\rho f_c}\left(|f| - \frac{(1-\rho)f_c}{2} - kB\right)\right)\right)$$

Table 6.3 Interference from globalstar to beidou-2

Modulation mode	Receiver bandwidth MHz	*SSC* dB/Hz	I_0 dBW/Hz	C/N_0 dB
MSK(4)	12.276	−72.06	−207.56	0.95
BPSK(4)	8.184	−72.48	−209.74	0.60
BPSK(8)	16.368	−72.48	−206.73	1.13
BOC(4,4)	16.368	−72.71	−206.96	1.08
MSK-BOC(4,4)	16.368	−72.08	−206.33	1.22
MSK-BOC(4,2)	12.276	−72.12	−207.62	0.94

Table 6.4 Maximum interference from beidou-2 to globalstar

Modulation mode	Receiver bandwidth MHz	*SSC* dB/Hz	I_0 dBW/Hz	C/N_0 95 %dB	C/N_0 5 %dB
MSK(4)	1.23	−67.28	−207.98	0.87	0.63
BPSK(4)	1.23	−66.43	−207.13	1.04	0.75
BPSK(8)	1.23	−69.33	−210.03	0.56	0.40
BOC(4,4)	1.23	−69.31	−210.01	0.56	0.40
MSK-BOC(4,4)	1.23	−69.03	−209.73	0.60	0.43
MSK-BOC(4,2)	1.23	−66.59	−207.29	1.00	0.73

Suppose the maximum number of Globalstar satellites in view of a Beidou-2 receiver has been fixed to 4 satellites at the same time. A typical noise PSD of −201.5 dBW/MHz has been considered. The intra-system interference of Beidou-2 is −222 dBW/MHz, while for Globalstar system under 95 % condition is −220 dBW/MHz and 5 % is −205 dBW/MHz [5]. Other assumptions are the same as the last paragraph. The simulation results are shown in Tables 6.3 and 6.4. Simulation results show that the effective carrier-to-noise loss caused by interference between Globalstar and Beidou-2 is quite little, which means that the interference between Globalstar and Beidou-2 is in a reasonable range, navigation services can be developed in S-band in future RNSS.

6.4 Signal Performance of S-band

Satellite navigation receiver's positioning, velocity measurement and timing function are all based on pseudo-range and carrier phase. However, pseudo-range is based on accurate code tracking, so code tracking precision is the key index to evaluate the receiver performance [3]. Of all the error sources that influence code tracking precision, thermal noise and multipath signal are the main error sources. Code tracking error and multipath error envelope reflect the receiver's code tracking precision influenced by thermal noise and multipath signal respectively, and are the key indexes to evaluate signal's code tracking performance. Gabor bandwidth which is signal's root mean square bandwidth decides the low bound of

signal's code tracking precision. Because it has nothing to do with the receiver's specific parameter, it is also the key index of evaluating satellite navigation signal. Here the Gabor bandwidth, code tracking error and multipath error envelope of the supposed modulations under certain signal parameter are analyzed.

Figure 6.1 is the curve of Gabor bandwidth varying with receiver bandwidth under different modulation. Gabor bandwidth is the weighted integral of signal power spectral density under the receiver bandwidth, when the weight is the square of the offset of the carrier frequency. Obviously, the power spectral of navigation signal has more high frequency component, the Gabor bandwidth is bigger and code tracking precision is higher. The bandwidth of S-band used for RNSS is 16.5 MHz, we can see that the Gabor bandwidth of MSK-BOC(4,2) and MSK-BOC(4,4) are bigger than others, which means the code tracking precision is higher.

Figure 6.2 is the curve of code tracking error varying with the carrier-to-noise under the condition of 1 Hz code tracking loop bandwidth, 16.5 MHz receiver bandwidth and correlator spacing of 0.1 chip intervals. Under the assumed conditions, the code tracking error curve of MSK-BOC(4,2) and MSK-BOC(4,4) is the lowest, which means the Cramer-Rao bound is lowest and the code tracking performance is best.

In view of the fact that in the real multipath environment the multipath to direct path signal amplitude ratio is lower than −6 dB, so the MDR value should be chosen as −10 dB when calculating the multipath error envelope. The estimation error is smaller while the MDR value is littler. Figure 6.3 is the multipath error

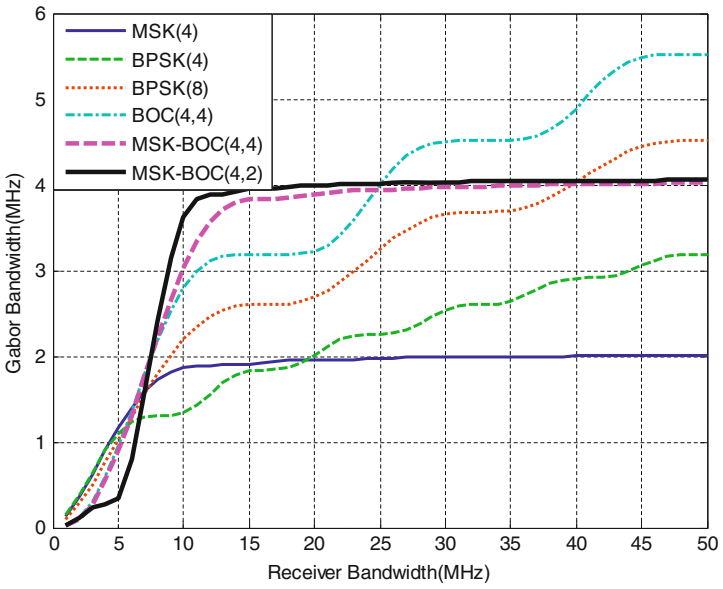


Fig. 6.1 *Curve* Gabor bandwidth varying with receiver bandwidth

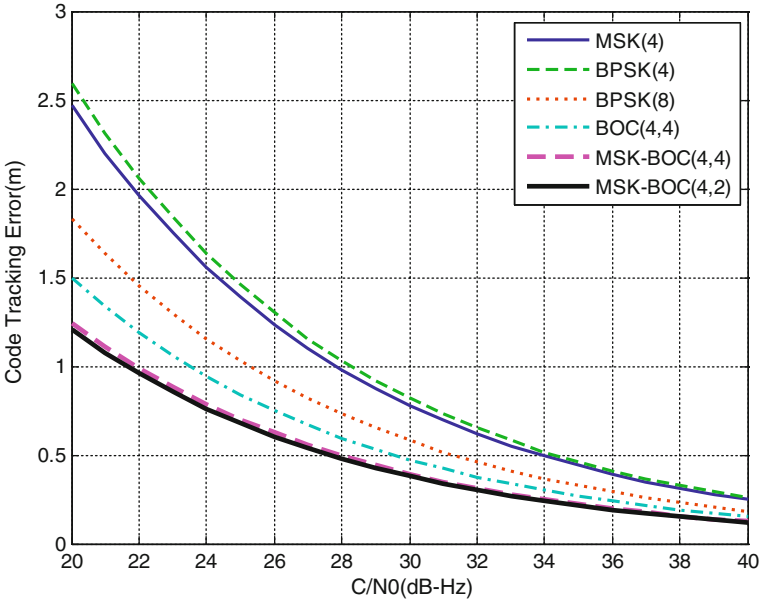


Fig. 6.2 *Curve* code tracking error varying with C/N0

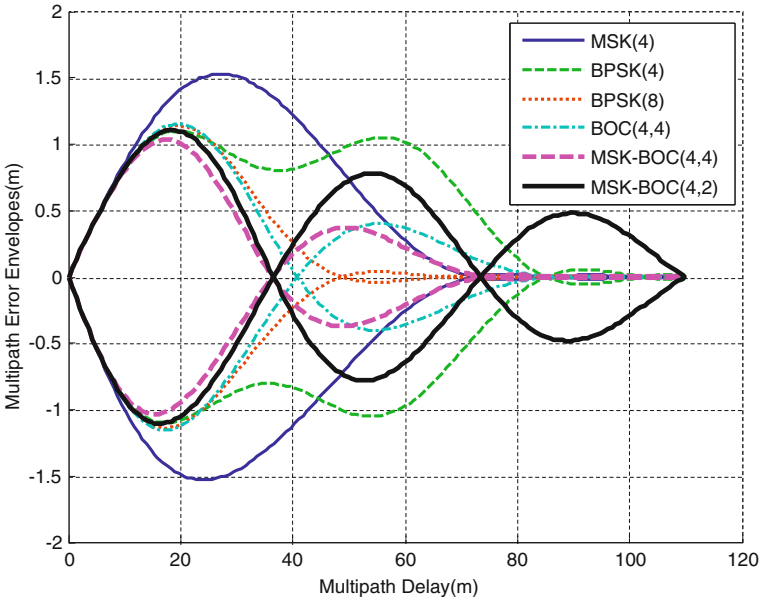


Fig. 6.3 *Curve* multipath error envelope

envelope when MDR is −10 dB, the receiver bandwidth is 16.5 MHz, the delay of multipath signal is one chip interval and 0.05 chip time of the correlator interval. The horizontal axis is the distance difference of reflection path and direct path and the vertical axis is the multipath error both units are all meter.

6.5 Conclusion

In this paper, the minimum shift keying—binary offset carrier (MSK-BOC) modulation was proposed to be used for Beidou-2 S-band. The compatibility between Beidou-2 and Beidou-1/Globalstar system was firstly evaluated, and then signal's Gabor bandwidth, code tracking error and multipath error envelope were analyzed. Results indicate that compared with the traditional BPSK and BOC modulation signals, using MSK-BOC(4,2) and MSK-BOC(4,4) modulation in S-band perform well in compatibility as well as tracking precision. The research will provide an instructive reference for Beidou-2 S-band's signal design and engineering implementation.

References

1. Mateu I, Boulanger C, Issler J et al (2009) Exploration of possible GNSS signals in S-band. In: GNSS ION 2009, Savannah
2. Xiulin H, Yihang R et al (2010) New option for modulation in GNSS signal design. Syst Eng Electron 32(9):P1962–P1967
3. Ran Y (2011) Study of GNSS signal modulation and frequency compatibility. Huazhong University of Science and Technology, Wuhan
4. Yang Z, Guangxia L et al (2011) Research on electromagnetic environment of satellite navigation in S-band. Geomatics Inf Sci Wuhan Univ 36(10):0
5. Mateu I, Paonni M, Issler J, et al. A search for spectrum: GNSS Signals in S-Band Part 2. In: Inside GNSS
6. Ran Y, Hu X, et al (2010) Compatibility analysis of compass navigation signals. China Satell Navig Conf 40(5):P676–684
7. Tang Z, Zhou H, Hu X, et al (2010) Research on performance evaluation of compass signal. China Satell Navig Conf 40(5):P592–602

Chapter 7
Application and Development of High Power Passive Device in Satellite System

Yanjiang Wei, Guang Yang and Kun Shan

Abstract The high power passive devices are playing a key role in the navigation satellite. Navigation is becoming an essential for our daily life as well as military applications. The GPS system, GLONASS system and Beidou system are all enhancing their anti-jamming capability by increasing more output power. Besides power amplifier and antenna, the power enhancing technology also calls for new demand to passive devices. High power capability, low loss, multipaction-free, low passive intermodulation (PIM), small size and weight are required. Typical high power passive devices are presented in this paper. Based on navigation application, the tendency of such devices is discussed.

Keywords Navigation satellite · High microwave power · Passive device

7.1 Introduction

Multipaction is also called secondary electron multipacting effect. It occurs when RF fields accelerate electrons in a vacuum and cause them to impact with a surface. If the frequency, field intensity and distance between metal surfaces of the RF meet certain range, which called falling into multifactor sensitive area and the coefficient of secondary electron yield from metal surfaces reaches greater than 1, these secondary electrons are accelerated by RF field and impact on metal surface, motivating more electrons. When the number of the increased electrons and escaped electrons strike a balance and produce electron clouds, the multipaction is happened.

Y. Wei (✉) · G. Yang · K. Shan
Southwest Institute of Electronic Equipment, Chengdu, China
e-mail: weiyj.cat@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_7,

The phenomenon of multipaction was first observed by Guttons in 1924. But confined to the space technology at that time and limited microwave power in space, people didn't know the mischief of multipaction profoundly. With the development of space technology, the application of microwave output power is much high. The multipaction becomes a vital constraining part of space technology development.

The navigation system all in European Galileo, American GPS, Chinese Beidou have the demand of hectowatt in L band [1]. In this case of high power, the danger of multipaction can not be ignored. With the increase of carrier wave quantity, different frequencies compose high peak power. Restraining multipaction has already become a key technology in satellite system.

7.2 Effect of Multipaction

7.2.1 Dielectric

Dielectric filling can decrease the mean free path of electrons and reduce the energy of electron impacting on metal surface. However, the secondary electron yield of dielectric is easier than metal surface and the dielectric materials release gases when heated, which result in gas discharge locally. In our experiment, we have observed multipaction of dielectric and metal surface several times. Therefore, we must use dielectric filling cautiously.

7.2.2 Surface Condition

There are mainly three factors of surface condition which effect multipaction. The first one is the surface coating material. Different material possesses different secondary electron emission coefficient. The secondary electron emission coefficient of gold and silver is less than common metal, so coating gold or silver on components' surface can acquire high threshold value of multipaction. Alodine surface treatment can achieve much high threshold value of multipaction.

Secondly smooth surface has great effect on the threshold value of multipaction. The aculeated part of device surface can assemble high electric field, which is a weak point in multipaction.

The third one is the cleanliness. The pollutant can greatly reduce the threshold value of multipaction.

7.2.3 Temperature

Temperature has two influences on multipaction. One is the surface state lying on temperature. High temperature leads to high energy of its own electron, which can get away from nucleus bound into secondary electron. Generally the breakdown voltage of multipactor is inversely proportional to the absolute temperature [2]. Therefore we need pay attention to heat radiation in high power application.

The other one is that low temperature can lead to dielectric contraction and make air gap, which may cause multipactor risk area, such as SMA joint.

7.3 High Power Passive Device

Spaceborne high power passive device is an important part in navigation satellite emission system. This paper introduces several kinds of high power passive device in navigation satellite.

7.3.1 Isolator

In navigation satellite, isolators mainly protect amplifiers in transmitters and improve their output standing wave ratio. With the increase of output power of navigation satellite, isolators need to bear more power from a few watts to one hectowatt. Currently the power has already been used to 150 W, and has a bullish trend now.

The power endurance capacity of space isolators depends on multipaction control and heat radiation. For the first problem of multipaction control, it's very difficult to prevent it using hectowatt power in 6 dB allowance because the navigation L band is sensitive to multipaction. We must take steps to solve this problem. The common method is choosing proper metal material and coating in central conductor of isolators. Usually we use beryllium bronze and the surface coating gold to decrease the secondary electron emission coefficient. Furthermore we increase the machining accuracy to avoid aculeated structure. For the other problem of heat radiation, there are two part of heat source in isolators. One part is the transmission loss in forward direction which converts into thermal energy causing temperature rise in isolators' resonant structures, including gyromagnetic substrate, central conductor, ground plane, permanent magnet. The other part is the reflected power which is absorbed by isolation resistance changing into heat energy. Current isolators mainly use ferrite devices. Their magnetic field is sensitive to the temperature. Therefore wide temperature design is needed. Effective measure is using earth permanent magnet, taking magnetic temperature compensation measures and employing ferrite substrate which changes less with

temperature. In addition isolators should be installed on heat sink in order to make sure that the ground plane can work in proper temperature range.

According to the proper use of navigation satellite, we have designed one L band isolator, which can solve the problem of multipaction and heat radiation well. At the 50 MHz working bandwidth and working temperature between −40 and +70 °C, the forward insertion loss is less than 0.3 dB, reverse loss is more than 20 dB, VSWR $\leq$ 1.2, average power capacity is 150 W. Its size is 72 × 62 × 22 mm. Figure 7.1 shows the test result of isolator.

At present, the power capacity of coaxial isolators in L band from abroad reaches 230 W. Figure 7.2 shows one isolator in Galileo satellite. This isolator combines two isolators into one by using a common isolation resistance. At the 50 MHz working bandwidth and working temperature between −40 and +70 °C, the forward insertion loss is less than 0.25 dB, reverse loss is more than 23 dB, VSWR $\leq$ 1.2, average power capacity is 230 W.

7.3.2 *Multiplexer*

Multiplexer is used for composing multi-channel power of different frequencies into one channel. The multipaction control of multiplexer is a traditional hard nut to crack. The reason is that multiplexer is strong resonant structure with high field intensity. Multi-channel composes into one output channel, so the output channel has the highest power and leads to more high peak power. The peak power can

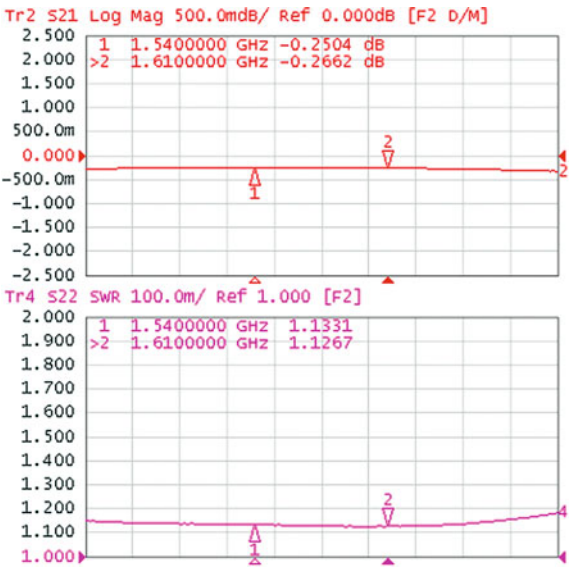


Fig. 7.1 Test result of isolator

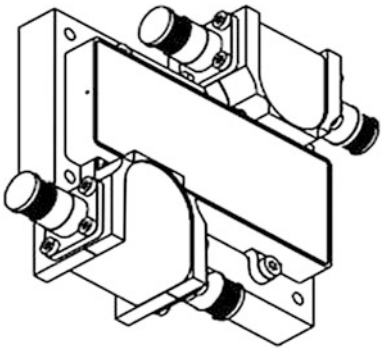


Fig. 7.2 L-band dual isolator

reach to N^2P in theory. For example three channels of 150 W can combine to 1,350 W of peak power (Fig. 7.3).

It is difficult to achieve 6 dB peak power allowance in designing. Besides, such power is hard to obtain when operating multipaction experiment. Actually it needs establishing time to arise multipaction [3]. Currently ESA has set up the standard of establishing multipaction whether or not by twenty electron transit times (T_{20}). Therefore the peak power is not important in preventing multipactor design. The envelope power higher than establishing time is even more important. This has close relationship with the frequency spacing of multitone signals.

Synthesis of envelope is only related to frequency spacing and nothing to do with carrier frequency with the definite carrier number and initial phase [4]. According to this characteristic, we make the simulation analysis below. Suppose three 150 W carrier with equal frequency spacing compose together, we analyze different frequency spacing effect on envelope. Figures 7.4, 7.5 show synthesis of envelope with 5 and 50 MHz frequency space respectively. P20 is regarded as equivalent power in multipaction effect, which satisfies threshold power in establishing time T20.

When the frequency space is 5 MHz, P20 is 1,209 W which is 8.1 times of single carrier power. And when the frequency space increases to 50 MHz, P20 is 394 W 2.6 times of single carrier power. From the Fig. 7.6 we can see P20 equivalent power and the scale factor of single carrier power.

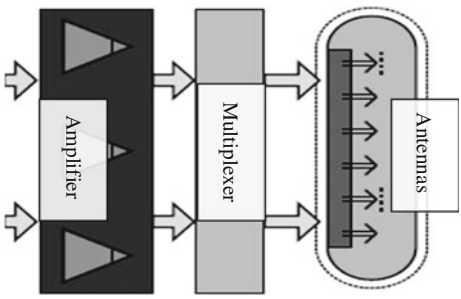


Fig. 7.3 OMUX diagram in system

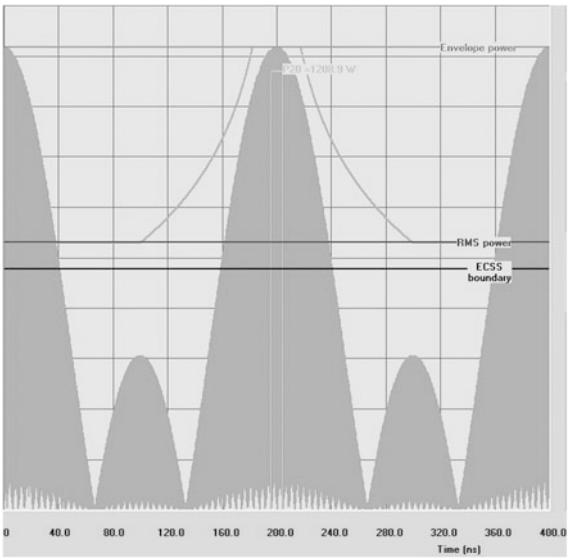


Fig. 7.4 Envelope of 5 MHz carrier spacing

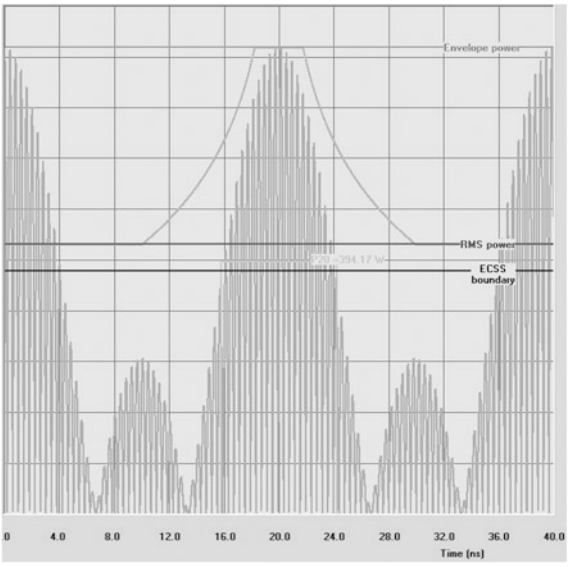


Fig. 7.5 Envelope of 50 MHz carrier spacing

It is thus clear that with increasing of the frequency spacing, P20 will decrease. When the frequency spacing raises up to 80 MHz, the P20 of three tone signal corresponds to singe carrier signal. We can come to a conclusion that, the peak superposition effect of pilot frequency is not obvious when the frequency space is large enough. So we use total average power to test. A research of ESA can provide evidence of experiment for the conclusion above [5].

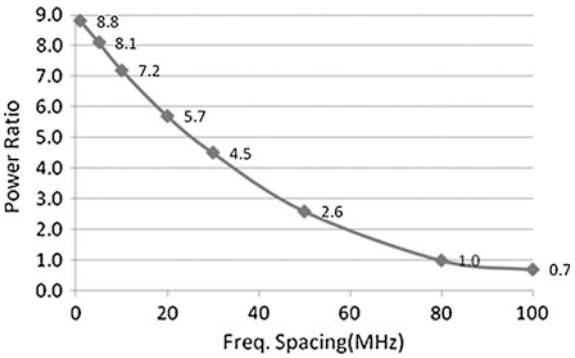


Fig. 7.6 Power ratio versus carrier spacing

We have already designed a triplexer with 150 W each channel according to navigation satellite application. Figure 7.7 shows our Triplexer. The threshold of multipaction is more than 6 dB each channel. The forward insertion loss is less than 0.4 dB.

7.3.3 High Power Switch

Single-pole double throw switch (SPDT) is usually used to switch main channel and backup channel in navigation satellite. The SPDT becomes vital device in switching main and backup channel for transmission sub system.

Currently the power of RF coaxial SPDT has already been used to 150 watts CW, and has a bullish trend now. It is different from normal SPDT, the switch in navigation satellite doesn't need switch frequently, it does change just when one channel in problem. Therefore we focus on the power capability and multipactor

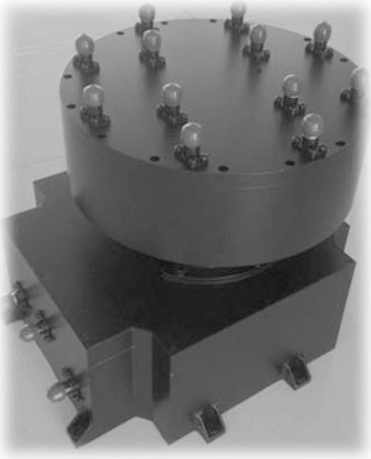


Fig. 7.7 Output triplexer

control when switches working. Figure 7.8 shows the difference between power capability and multipactor control ability of high power SPDT. From the curve we can see that, the lower frequency has stronger power capacity and weaker multipactor control ability, and vice versa.

We have designed a coaxial high power SPDT in L band, which can solve the problem of multipaction and power capacity well. At the 500 MHz frequency range and working temperature between −40 and +70 °C, the forward insertion loss is less than 0.2 dB, isolation is more than 70 dB, VSWR $\leq$ 1.2, average power capacity is 150 W and satisfy the 6 dB threshold in multipaction. Its size is 83 × 60 × 32.5 mm (Figs. 7.9, 7.10).

7.4 Future Developing Trends

With the increase of channel power of navigation satellite, we can foresee the future trend of development in high power passive devices as below.

1. High power capability
2. High multipactor control ability
3. Depressing of passive intermodulation.

In order to increase the EIRP of navigation satellite, one method is improving the antenna gain. The other hand is by using high output power amplifier, such as TWTA and SSPA. The project of Galileo has already done research in 200 CW

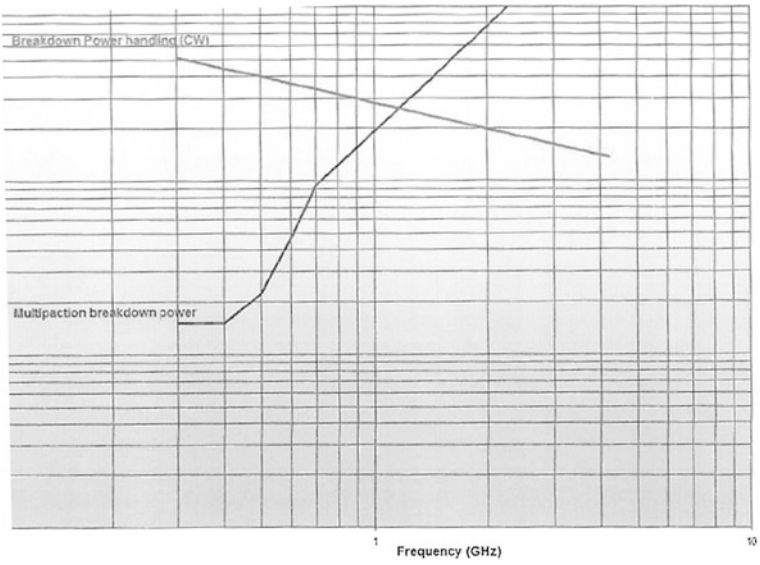


Fig. 7.8 Different of power handing and multipaction

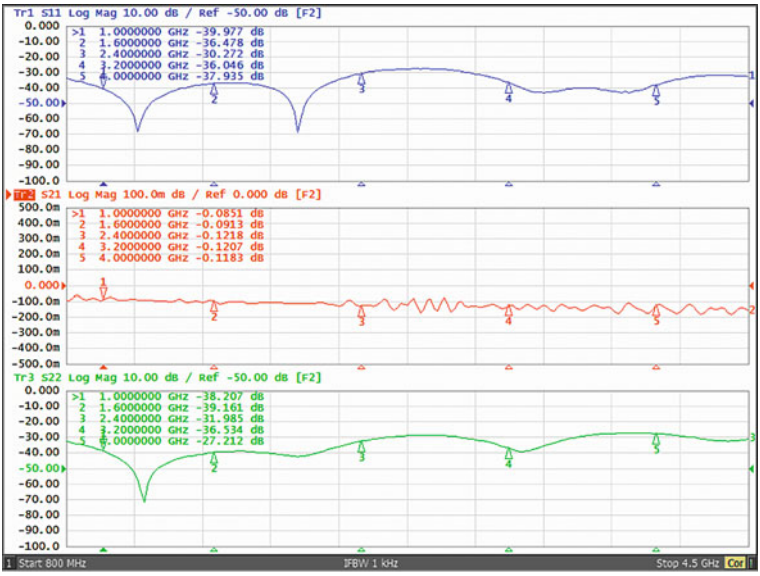
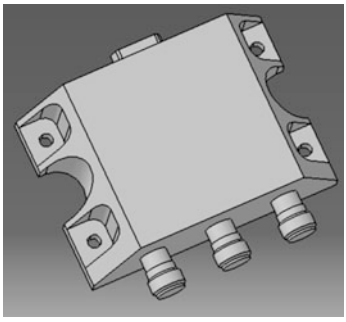


Fig. 7.9 Test result of SPDT

Fig. 7.10 High power SPDT



watts TWTA. The high power passive devices with 200 W larger power capacity are needed correspondingly. It's very difficult to design the passive devices both have high power output and multipaction restraining. We have already designed coaxial high power passive devices with 150 W average power and 6 dB multipactor threshold in navigation application. That is 600 W in peak power, which is the top level of domestic. However, this power level cannot satisfy future need. We should develop much high power passive devices not limited to coaxial structure.

The passive intermodulation is also important in passive devices. It is happened in multi-carrier system, due to its nonlinear characteristic result in different carrier of harmonic and difference frequency signals. These intermodulation will grow bigger with input power increases, then affect both the navigation signal quality

and receive channel. The mechanism of intermodulation is very complicated. From the current situation, the main measures taken are electroplating well in device surface, avoiding contact with different materials, careful use of magnetic material, whole surface cleaning in purification workshop.

Future development of navigation satellite will be in increasingly demand of high power passive components. In order to guarantee the performance of high power passive devices and high reliability, in addition to design more useful devices effectively, but also strictly control in the material, process, assembly, testing etc. The high power passive devices should be assembled in purification workshop and designed to ensure multipactor threshold margin of at least 6 dB. Besides, they should take multipactor discharge test before go into space to ensure the acceptance level of minimum 3 dB allowance.

Acknowledgments Sincere thanks to teacher Guan Zhaohui in the guidance and help in writing this paper.

References

1. Guang Y, Gang Q (2011) Development of space SSPA based on domestic electrical devices. Chin Satell Navig Conf 5:999–1002
2. Xuda W, Ying Z (2000) Multipaction of coaxial filter. Space Electron Technol 4:6–9
3. Xin Y, Junye C (2004) Analysis of establishing time on multipaction. Space Electron Technol 1:27–32
4. Xin Y, Junye C (2002) Analysis of multicarriers anti-multipation design. Space Electron Technol 4:42–47
5. Jianrong C, Xuda W (1999) Analysis of multipaction in satellite equipment. Space Electron Technol 1:19–23

Chapter 8
Analysis for Cross Correlation in Multiplexing

Yingxiang Liu, Xiaomei Tang, Rui Ge and Feixue Wang

Abstract Several multiplexing techniques may be used in next generation GNSS, including QPSK, Interplex, AltBOC, TD-AltBOC, Dual QPSK, POCET. Multiple signals on the same carrier could be multiplexed into a constant-envelope signal with these techniques. Increasing of signals would deteriorate the crossing correlation and its impact on receiving is analyzed. It is shown that cross correlation between signals in-phase has considerable influence; in contrast, the effect could be neglected to signals quadrature. Simulation and real data test is given in the end to support the conclusion, which could guide the selection of multiplex technique and spreading code of Compass.

Keywords GNSS · Multiplexing · Cross correlation · PRN code selection

8.1 Introduction

With diversification of global navigation satellite system (GNSS) application, navigation signals become further complicated. The next generation GNSS transmit multiple signals on a single carrier to meet the demand of different users.

GNSS signals on the same carrier should be multiplexed into a constant-envelope signal [1]. Variable multiplexing techniques may be used in next generation GNSS, which includes quadrature phase shift keying (QPSK) [2], Interplex [3], alternative binary offset carrier (AltBOC) [4], time division AltBOC (TD-AltBOC) [5], Dual QPSK [6], and phase-optimized constant-envelope transmission (POCET) [7].

Y. Liu (✉) · X. Tang · R. Ge · F. Wang
Satellite Navigation and Positioning R&D Center, School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073 Hunan, China
e-mail: inshine1986@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_8,

For traditional GNSS receivers, only cross correlation between signals of different satellites need to be considered. As the relative motion of different satellites is time-variant, the impact of cross correlation could be neglected when there is no big difference in signal power. In next generation GNSS, multiple civil signals are modulated on a single carrier, and their phase and motion are always the same; therefore the cross correlation is time-invariant, which leads to unchangeable tracking bias. However, there is no analysis on this issue yet.

At the beginning of the paper, a simplified model from the aspect of signal receiving is introduced, and then the impact of cross correlation on signal tracking is theoretically analysed. In the end, simulation and real data test is given to support the results. The conclusion can guide the signal design and pseudorandom noise (PRN) code selection of Compass.

8.2 Simplified Model of Multiplexed Signal

Multiplexing techniques have been and would be used in GNSS, which includes QPSK, Interplex, AltBOC, TD-AltBOC, Dual QPSK, and POCET. Theoretical expressions of some techniques are quite complicated because of intermodulation components [8]. However, from the aspect of signal receiving, all of these multiplexed signals could be modeled as follows:

$$r(t) = \sum_{i=1}^{K} d_i(t)c_i(t)\cos(2\pi f_i t + \theta_i) + IM(t) + n(t) \tag{8.1}$$

where

K	is the number of useful signal components;
$d_i(t)$ and $c_i(t)$	is navigation message and ranging code of signal i, which is modulated on the carrier of frequency f_i and initial carrier phase θ_i;
$IM(t)$	is the intermodulation components for the purpose of constant-envelope multiplexing;
$n(t)$	represents the noise, which would be neglected because it has no effect on the mean value of tracking results.

Next, these variables would be explained in details, and some would be confined for the simplification of analysis.

8.2.1 Number of Signal Components K

Two signal components is the simplest condition of multiplexing, and the impact of cross correlation of multiple signal components can be easily concluded

according to that of two signal components. In order to simplify the analysis, the number of signal components K is confined to 2.

8.2.2 Ranging Code c(t)

Ranging code $c(t)$ has different meanings for different modulations, such as:

1. For binary phase shift keying (BPSK) modulation, $c(t)$ represents raw PRN code;
2. For binary offset carrier (BOC), composite BOC (CBOC) and time-multiplexed BOC (TMBOC) modulation, $c(t)$ represents PRN code modulated with corresponding subcarrier;
3. For time division data multiplexing (TDDM) modulation, $c(t)$ represents PRN code with three levels of ± 1 and 0, and its waveform is illustrated in Fig. 8.1.

8.2.3 Carrier Frequency f

Multiple signal components can be multiplexed on a single carrier; however, it does not mean that the all of the signal components have the same carrier frequency. So the variable f_i in (8.1) may not be the same, such as the upper and lower band of AltBOC and TD-AltBOC modulated signals, which have different carrier frequency.

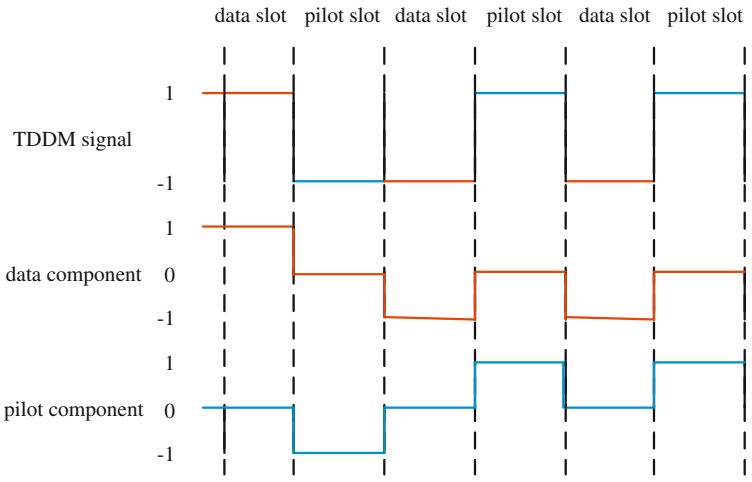


Fig. 8.1 Waveform of data/pilot component of TDDM signal

When the carrier frequency of two signals is not identical, the cross correlation could be neglected because the unwanted signal component could be filtered. So only the cross correlation between signals with the same carrier frequency is analyzed in the paper.

8.2.4 Intermodulation Component IM(t)

The correlation between the intermodulation component $IM(t)$ and any signal component should be zero. If not, the correlation peak would be asymmetric. So for a well-designed GNSS signal, the intermodulation component could be neglected for signal receiving.

8.2.5 Simplified Model

Based on the above analysis, the received signal can be simplified as

$$r(t) = d_1(t)c_1(t)\cos(2\pi f_c t) + d_2(t)c_2(t)\cos(2\pi f_c t + \varphi) \tag{8.2}$$

where
f_c is the carrier frequency;
φ is the relative phase between two carriers.

The relative phase of different multiplexing techniques can be categorized as three cases:

1. $\varphi = 0$, as signals of the same sideband in TD-AltBOC modulation;
2. $\varphi = \pi/4$, as signals in Dual QPSK modulation;
3. $\varphi = \pi/2$, as signals of the same sideband in AltBOC modulation.

Next, the impact of cross correlation on signal $c_1(t)$ tracking is analyzed based on the simplified model.

8.3 Impact of Cross Correlation

The discriminator output of code and carrier tracking loop in consideration of cross correlation is given at first, and then the theoretical expression of code and carrier tracking bias caused by cross correlation is derived.

8.3.1 Model of Loop Discriminator

The carrier frequency of received signals is assumed to have been accurately estimated. Two orthogonal carriers $2\cos(2\pi f_c - \theta)$ and $-2\sin(2\pi f_c - \theta)$ are locally generated to mix with the received signal, and the baseband signal after filtering is

$$\begin{aligned} s(t) &= LP\{r(t) \times 2\exp(-j2\pi f_c t + \theta)\} \\ &= d_1(t)c_1(t)e^{j\theta} + d_2(t)c_2(t)e^{j(\theta+\varphi)} \end{aligned} \tag{8.3}$$

where $LP\{\cdot\}$ represents low pass filtering.

Assuming that the code phase delay is τ and the early-late space is 2δ, the correlation values of early, prompt and late branch are

$$\begin{aligned} SE_k &= \frac{1}{T_c}\int_{t=kT_c}^{(k+1)T_c} s(t)c_1(t+\tau+\delta)dt \\ &= d_{1k}R_{11}(\tau+\delta)e^{j\theta} + d_{2k}R_{12}(\tau+\delta)e^{j(\theta+\varphi)} \end{aligned} \tag{8.4}$$

$$\begin{aligned} SP_k &= \frac{1}{T_c}\int_{t=kT_c}^{(k+1)T_c} s(t)c_1(t+\tau)dt \\ &= d_{1k}R_{11}(\tau)e^{j\theta} + d_{2k}R_{12}(\tau)e^{j(\theta+\varphi)} \end{aligned} \tag{8.5}$$

$$\begin{aligned} SL_k &= \frac{1}{T_c}\int_{t=kT_c}^{(k+1)T_c} s(t)c_1(t+\tau-\delta)dt \\ &= d_{1k}R_{11}(\tau-\delta)e^{j\theta} + d_{2k}R_{12}(\tau-\delta)e^{j(\theta+\varphi)} \end{aligned} \tag{8.6}$$

where

SE, SP, SL	is the correlation values of the early, prompt and late branch;
T_c	is the coherent integration time;
$R_{11}(\tau)$	is the auto-correlation function of $c_1(t)$;
$R_{12}(\tau)$	is the cross correlation function between $c_1(t)$ and $c_2(t)$

Delay locking loop (DLL) estimates the code phase delay according to the correlation values of the early and late branch. Noncoherent early minus late power (EMLP) discriminator is usually used in DLL and its expression is [9]

$$\begin{aligned} \varepsilon_D &= |SE_k|^2 - |SL_k|^2 \\ &= R_{11}^2(\tau+\delta) - R_{11}^2(\tau-\delta) + R_{12}^2(\tau+\delta) - R_{12}^2(\tau-\delta) \\ &\quad + 2d_{1k}d_{2k}R_{11}(\tau+\delta)[R_{12}(\tau+\delta) - R_{12}(\tau-\delta)]\cos\varphi \end{aligned} \tag{8.7}$$

Phase locking loop (PLL) track the carrier phase according to correlation values of prompt branch. Costas discriminator is usually used to make PLL insensitive to sign change [2]. The two-quadrant inverse tangent discriminator is

$$\begin{aligned}\varepsilon_P &= \operatorname{atan}\left(\frac{\operatorname{Im}\{SP_k\}}{\operatorname{Re}\{SP_k\}}\right) \\ &= \operatorname{a\,tan}\left[\frac{d_{1k}R_{11}(\tau)\sin\theta + d_{2k}R_{12}(\tau)\sin(\theta+\varphi)}{d_{1k}R_{11}(\tau)\cos\theta + d_{2k}R_{12}(\tau)\cos(\theta+\varphi)}\right]\end{aligned} \tag{8.8}$$

It can be seen that if cross correlation could be neglected compared with auto-correlation, PLL and DLL can estimate the carrier phase and code phase without any bias.

8.3.2 Code and Carrier Tracking Bias

The code loop track stably when the code phase delay τ_0 satisfy the following equation:

$$\varepsilon_D = 0|_{\tau=\tau_0} \tag{8.9}$$

As the cross correlation function has no explicit expression, there is no closed solution for code phase delay τ_0, and its approximate value would be analyzed. The DLL discriminator can be considered as the sum of auto-correlation term ε_a and cross correlation term ε_c:

$$\varepsilon_a \triangleq R_{11}^2(\tau+\delta) - R_{11}^2(\tau-\delta) \tag{8.10}$$

$$\begin{aligned}\varepsilon_c \triangleq\ & R_{12}^2(\tau+\delta) - R_{12}^2(\tau-\delta) \\ & + 2d_1d_2R_{11}(\tau+\delta)[R_{12}(\tau+\delta) - R_{12}(\tau-\delta)]\cos\varphi\end{aligned} \tag{8.11}$$

When the relative phase φ equals $\pi/2$, the cross correlation term is

$$\varepsilon_c = R_{12}^2(\tau+\delta) - R_{12}^2(\tau-\delta) \tag{8.12}$$

As $\tau_0 \ll \delta$, the approximate value of code tracking bias can be obtained by

$$\tau_0 = \frac{R_{12}^2(\delta) - R_{12}^2(-\delta)}{4KR_{11}(\delta)} \tag{8.13}$$

where K is the gradient of the auto-correlation function at $\tau = \delta$.

From (8.13), when φ equals $\pi/2$, the code tracking bias is related to auto-correlation, cross correlation and the early-late space, so the tracking bias caused by cross correlation is a fixed value.

Take the narrow correlators for BPSK signal with 1 MHz code rate as an example, assuming the peak-to-peak value of cross correlation is no more than 1/100, then the tracking bias τ_0 is less than 1 cm. So when the carrier phase of two PRN codes is orthogonal, the tracking bias caused by cross correlation is constant and negligible.

When the relative phase φ equals to 0 or $\pi/4$, the cross correlation term in the discriminator can be approximated as

$$\varepsilon_c \approx 2\cos\varphi d_{1k}d_{2k}R_{11}(\delta)[R_{12}(\delta) - R_{12}(-\delta)] \tag{8.14}$$

Similarly, the code tracking bias under this situation can be obtained:

$$\tau_0 = \frac{2\cos\varphi d_{1k}d_{2k}[R_{12}(\delta) - R_{12}(-\delta)]}{4K} \tag{8.15}$$

By similar calculation under the same assumption, τ_0 can reach to 1.5 m on this condition. Thus, when the carrier phase of two PRN code is not orthogonal, the tracking bias is considerable. Besides, it can be seen from (8.15) that if $\varphi \neq \pi/2$ the code tracking bias is also related with the navigation message. As in real environment, the navigation message has periodicity within in a time span of hours, cross correlation would lead to periodical fluctuation in tracking bias.

The carrier loop tracking stably when the carrier phase delay θ_0 satisfy the following equation:

$$d_{1k}R_{11}(\tau_0)\sin\theta_0 + d_{2k}R_{12}(\tau_0)\sin(\theta_0 + \varphi) = 0 \tag{8.16}$$

According to (8.16), the approximate value of θ_0 can be obtained:

$$\theta_0 \approx -\frac{d_{1k}}{d_{2k}}\tan^{-1}\left[\frac{R_{11}(\tau_0)\sin(\varphi)}{R_{12}(\tau_0)}\right] \tag{8.17}$$

Under the same assumption that the peak-to-peak value of cross correlation is no more than 1/100, the carrier phase delay θ_0 is smaller than 0.57°, which can be ignored. Specifically, when two signal components are in-phase, the cross correlation would not cause any carrier tracking bias.

From the above analysis, the relative relationship between auto-correlation and cross correlation is irrelevant to satellite motion, and the noise would not influence the mean value of tracking bias, therefore the conclusion can also be applied in real environment.

8.4 Simulations and Real Data Testing

8.4.1 Simulation Results

The conclusion is verified by Monte Carlo simulations herein. TD-AltBOC and AltBOC signals are simulated, which corresponds to the case of $\varphi = 0$ and $\varphi = \pi/2$. The simulation parameters are listed in Table 8.1.

Only signal $c_1(t)$ is received in simulation, and simulation results are shown in Figs. 8.2 and 8.3.

Table 8.1 Simulation parameters

Parameters	Values
Multiplexing	AltBOC, TD-AltBOC
PRN Code	Lower sideband: c_1:PRN1, c_2:PRN5
	Upper sideband: c_3:PRN2, c_4:PRN3
Navigation message period	6 s
C/N_0	50dBHz
Coherent integration time	1 ms

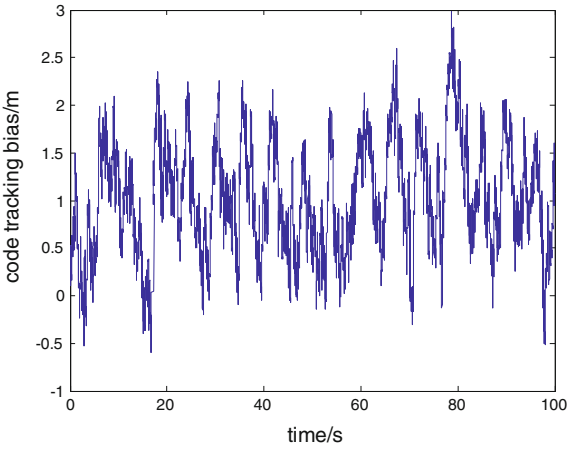


Fig. 8.2 Signal $c_1(t)$ code tracking bias of TD-AltBOC modulation

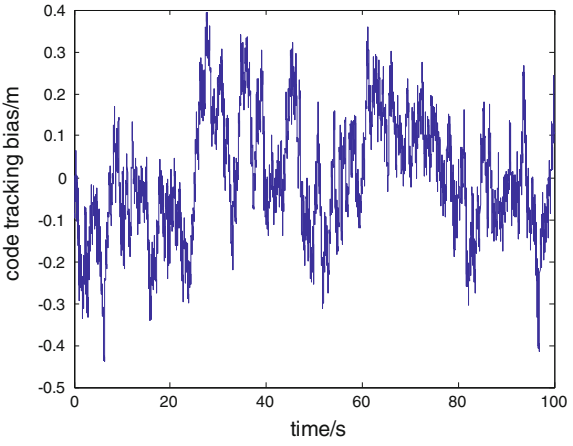


Fig. 8.3 Signal $c_1(t)$ code tracking bias of AltBOC modulation

It can be seen from the above results that cross correlation of TD-AltBOC caused much more significant tracking bias than that of AltBOC, and the tracking bias fluctuate with a period the same as the navigation message. The phenomenon is exactly consistent with the above theoretical analysis.

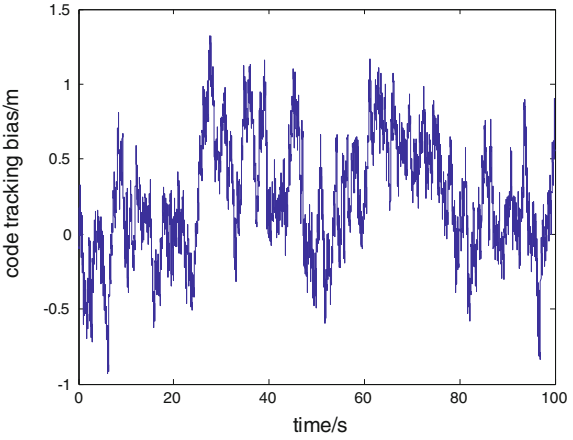


Fig. 8.4 Code tracking bias of $c_1(t)$ of TD-AltBOC modulation after replacing signal $c_2(t)$ PRN code with PRN5

Next, a comparative simulation by replacing PRN5 with PRN2 is carried out, and the result is illustrated in Fig. 8.4.

The simulation result shows that PRN code with better cross correlation performance can obviously reduce the tracking bias.

8.4.1.1 Real Data Testing

Due to the limitation of current GNSS signal, real signals from space cannot be used to prove the conclusion. An equivalent verification could be made by changing the constellation of GPS simulator. By setting each satellite as GEO satellite, the code phase delay between different satellites is time invariant, so it is equivalent as one satellite transmitting multiple signals. The result is shown in Fig. 8.5.

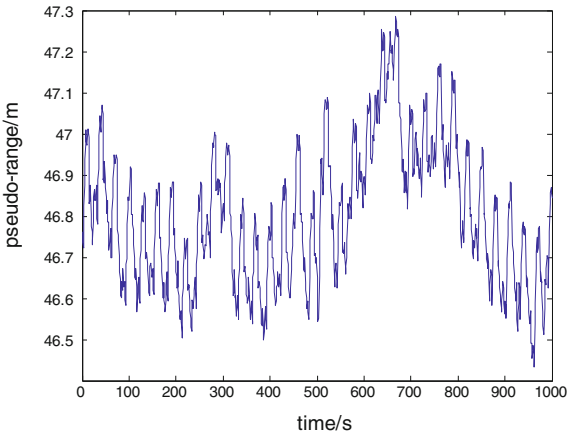


Fig. 8.5 Pseudo-range of GPS PRN1 satellite

From the above result, the pseudo-range has a fluctuation with 30 s period, which is the same as the GPS massage frame structure. So the conclusion is verified by real data testing.

8.4.2 Conclusions

The impact of cross correlation in next generation GNSS is analyzed in this paper. It is concluded that when two signal components are not orthogonal, the cross correlation would cause large code tracking bias with periodic fluctuation.

Therefore, the AltBOC modulation has more advantage than TD-AltBOC in reducing the impact of cross correlation. If two signal components are not non-orthogonal, both PRN code should be specifically selected so that the cross correlation within ± 1 code chip can be neglected. In multiplexing technique and PRN code selection of Compass, special attention should be paid on the issue.

References

1. Avila-Rodriguez JA (2008) On generalized signal waveforms for satellite navigation. PhD Thesis, Faculty of Aerospace Engineering, University FAF Munich, Munich, Germany
2. Kaplan ED, Hegarty CJ (2006) Understanding GPS principles and applications, 2nd edn. Publishing House of Electronics Industry, Beijing, pp 88–89
3. Butman S, Timor U (1972) Interplex-an efficient multichannel PSK/PM telemetry system. IEEE Trans Commun 20(3):415–419
4. Lestarquit L, Artaud G, Issler J (2008) AltBOC for dummies or everything you always wanted to know about AltBOC. In: Proceedings of the 21st international technical meeting of the satellite division of the institute of navigation, Savannah, 16–19 Sept 2008, pp 961–970
5. Tang Z (2012) TD-AltBOC: A New COMPASS B2 Modulation, CSNC 2012
6. Zhang K, A signal generation method of Dual QPSK. Patent
7. Dafesh PA, Cahn CR (2009) Phase-optimized constant-envelope transmission (POCET) modulation method for GNSS signals. In: Proceedings of the 22nd international technical meeting of the satellite division of the institute of navigation, Savannah, 22–25 Sept 2009, pp 2860–2866
8. Zhang K, Analytical model of POCET for COMPASS B1/B3 signals, GNSS 2012
9. Van Dierendonck AJ (1996) GPS receivers. In: Global Positioning System: Theory and Applications, New York: AIAA, 329—408

Chapter 9
Performance Analysis on Single Sideband of TD-AltBOC Modulation Signal

Tao Yan, Jiaolong Wei, Zuping Tang and Zhihui Zhou

Abstract TD-AltBOC modulation signal is a new signal. Similar to AltBOC(15,10) signal, TD-AltBOC signal contains lower and upper sideband signals. From the view of the selection of PN codes for lower and upper sideband signals' pilot components, there are two schemes to achieve TD-AltBOC, namely to choose different PN codes or the same PN code. For considering the cost and complexity of receivers, only the lower sideband signal may be received for mass market receivers. When the receiver is only tracking lower sideband signal's pilot component, due to the non-ideal PN codes and the effect of upper sideband signals, a asymmetrical correlation function may be obtained, which causes the cross zero point of discriminator curve to move, so code tracking bias error is introduced. In this paper, we analyze the reason and condition why correlation function is asymmetry, and make performance analysis on single sideband of TD-AltBOC modulation Signal. Simulation results show that the selection of PN codes for data component may cause correlation function asymmetry. In order to reduce the effect on single sideband reception and reach the same performance with AltBOC(15,10), the selection of PN codes has to meet requirement of this paper.

Keywords TD-AltBOC · PN code · Code tracking bias error

T. Yan (✉) · J. Wei (✉) · Z. Tang · Z. Zhou
The Department of Electronic and Information Engineering of Huazhong University of Science and Technology, Wuhan, China
e-mail: iamyantao@126.com

J. Wei
e-mail: jlwei@mail.hust.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_9,

9.1 Introduction

The constant envelope AltBOC(15,10) modulation type is chosen as Galileo E5 signal's modulation type [1, 2]. AltBOC(15,10) signal is a wideband signal, whose lower and upper sideband can provide different services. This modulation type not only can meet the requirement of constant envelope signal for high power amplifier, but also can reduce signal's implementation complexity, avoiding the excessive demand on the filter [3]. And AltBOC(15,10) modulation signal is also a candidate at B2 frequency for China's Beidou global navigation satellite system [4]. In Ref. [5] the new TD-AltBOC(15,10) modulation, which can be used as B2 signal's implementation method, is proposed. Different form the constant envelope AltBOC(15,10) modulation for Galileo E5 signal, TD-AltBOC(15,10) modulation signal uses time division technique, that means data and pilot component at lower and upper sideband are multiplexed together by time division manner, respectively. The ranging and anti-multipath performance of TD-AltBOC(15,10) signal is similar to Galileo E5 signal's. And Interoperability with Galileo AltBOC signal can be easily achieved. Specially, the choose of PN code for pilot component at lower and upper sideband has a significant impact on generation and receiving methods. So there are two kinds of implementation scheme for TD-AltBOC signal. One is to use the same PN codes for pilot component at lower and upper sideband; the other is to use the different PN codes.

Similar to Galileo E5 signal, TD-AltBOC(15,10) modulation signal is also a wideband signal, which can achieve potential centimetre level ranging accuracy. But for considering the cost and complexity of receivers, only the lower sideband signal may be received for mass market receivers. However, when the receiver is only tracking lower sideband signal's pilot component, due to the non-ideal PN codes and the effect of upper sideband signals, a asymmetrical correlation function may be obtained, which causes the cross zero point of discriminator curve to move, so code tracking bias error is introduced. In this paper, we analyze the reason and condition why correlation function is asymmetry, and make performance analysis on single sideband of TD-AltBOC modulation Signal. Simulation results show that the selection of PN codes for data component is the main reason that correlation function is asymmetry, which is not related to the selection of PN codes for upper sideband signals' pilot component. In order to reduce the effect on single sideband reception and achieve the same performance with Galileo AltBOC(15,10), the selection of PN codes has to meet requirement of this paper.

The rest of the paper is organized as follows. The TD-AltBOC modulation signal is introduced in Sect. 9.2. Then the problem that cross zero point of discriminator curve may move when the receiver is only tracking lower sideband signal's pilot component is given out, and the reason is analyzed in Sect. 9.3. In Sect. 9.4, performance simulation results of single sideband of TD-AltBOC modulation Signal are presented. In Sect. 9.5, we conclude some conclusions.

9.2 TD-AltBOC Modulation Signal

TD-AltBOC(m,n) modulation signal is defined by the PN code rate R_c = n*1.023 MHz and subcarrier frequency f_s = m*1.023 MHz. The mathematical expression of the TD-AltBOC baseband signal is [5]:

$$\begin{aligned} s(t) = &\left[d_A(t)c_{AD}(t) + c_{AP}(t)\right]\left[SC_{B,\cos}(t) - jSC_{B,\sin}(t)\right] \\ &+ \left[d_B(t)c_{BD}(t) + c_{BP}(t)\right]\left[SC_{B,\cos}(t) + jSC_{B,\sin}(t)\right] \end{aligned} \tag{9.1}$$

where d_A(t) is navigation message bit waveform of lower sideband data channel, d_B(t) is navigation message bit waveform of upper sideband data channel, c_{AD}(t) is PN code waveform of lower sideband data channel, c_{AP}(t) is PN code waveform of lower sideband pilot channel, c_{BD}(t) is PN code waveform of upper sideband data channel, c_{BP}(t) is PN code waveform of upper sideband pilot channel. Their definition is:

$$\begin{aligned} c_{AD}(t) &= \sum_{l=-\infty}^{\infty}\sum_{k=0}^{N_{AD}-1} C_{AD}(k)p(t-(2N_{AD}l+2k)T_c) \\ c_{AP}(t) &= \sum_{l=-\infty}^{\infty}\sum_{k=0}^{N_{AP}-1} C_{AP}(k)p(t-(2N_{AP}l+2k+1)T_c) \\ c_{BD}(t) &= \sum_{l=-\infty}^{\infty}\sum_{k=0}^{N_{BD}-1} C_{BD}(k)p(t-(2N_{BD}l+2k)T_c) \\ c_{BP}(t) &= \sum_{l=-\infty}^{\infty}\sum_{k=0}^{N_{BP}-1} C_{BP}(k)p(t-(2N_{BP}+2k+1)T_c) \end{aligned} \tag{9.2}$$

where C_{AD} is the PN code sequence of lower sideband data channel, whose length is N_{AD}. C_{AP} is the PN code sequence of lower sideband pilot channel, whose length is N_{AP}. C_{BD} is the PN code sequence of upper sideband data channel, whose length is N_{BD}. C_{BP} is the PN code sequence of upper sideband pilot channel, whose length is N_{BP}. $T_c = 1/R_c$ is the chip period. $p(t)$ is defined as:

$$p(t) = \begin{cases} 1 & 0 \le t < T_c \\ 0 & others \end{cases} \tag{9.3}$$

$SC_{B,cos}(t)$ and $SC_{B,sin}(t)$ are the sine phase and cosine phase square wave subcarrier, respectively. That is:

$$\begin{aligned} SC_{B,\cos}(t) &= \mathrm{sign}(\cos(2\pi f_s t)) \\ SC_{B,\sin}(t) &= \mathrm{sign}(\sin(2\pi f_s t)) \end{aligned} \tag{9.4}$$

From the above, we can know that data component and pilot component at lower and upper sideband are multiplexed by time division method. This means that only the data component is transmitted at even time slot, and only the pilot component is transmitted at odd time slot. When C_{AP} is the same with C_{BP}, the baseband signal's expression at even time slot can be simplified as:

$$s(t) = 2 \cdot c_{AP}(t) \cdot SC_{B,\cos}(t) \tag{9.5}$$

So if the PN codes of pilot component at lower and upper sideband are the same, the generation and receiving methods of TD-AltBOC modulation signal could be simplified.

9.3 Performance Analysis

Firstly we get the approximate baseband expression of TD-AltBOC signal. Because the complex subcarrier is a periodic signal, it can be expressed as Fourier series expansion. That is:

$$\begin{aligned} SC_{B,\cos}(t) - jSC_{B,\sin}(t) &\approx \tfrac{4}{\pi} e^{-j2\pi f_s t} + \tfrac{4}{3\pi} e^{j2\pi \cdot 3 f_s t} \\ SC_{B,\cos}(t) + jSC_{B,\sin}(t) &\approx \tfrac{4}{\pi} e^{j2\pi f_s t} + \tfrac{4}{3\pi} e^{-j2\pi \cdot 3 f_s t} \end{aligned} \tag{9.6}$$

Let P_r is the total received power of TD-AltBOC signal. When only the lower sideband signal is received, the received power of desired signal is not $P_r/2$, but $(4/\pi)^2/2/P_r/2 = 0.8106^* P_r/2$. So from the view of single sideband signal, the combination loss of TD-AltBOC modulation is 81.06 %.

Let $S_{AD}(t) = d_A(t)c_{AD}(t)$, $S_{BD}(t) = d_B(t)c_{BD}(t)$, $S_{AP}(t) = c_{AP}(t)$, $S_{BP}(t) = c_{BP}(t)$, P_T is the transmitted power of TD-AltBOC signal from single satellite. The transmitted signal is approximately expressed as:

$$\begin{aligned} s_T(t) \approx & \frac{\sqrt{P_T}}{2}\left[s_{AD}(t) + s_{AP}(t)\right]\left[\frac{4}{\pi} e^{-j2\pi f_s t} + \frac{4}{3\pi} e^{j2\pi \cdot 3 f_s t}\right] \\ & + \frac{\sqrt{P_T}}{2}\left[s_{BD}(t) + s_{BP}(t)\right]\left[\frac{4}{\pi} e^{j2\pi f_s t} + \frac{4}{3\pi} e^{-j2\pi \cdot 3 f_s t}\right] \end{aligned} \tag{9.7}$$

Assumed the channel's transfer function from satellite to the receiver is $H(f)$. the total received power of TD-AltBOC signal is P_r. Only the pilot component at lower sideband is received. The signal's received bandwidth is β_r. The received signal is:

$$\begin{aligned} & s_r(t) = s_{rs}(t) \otimes h(t), h(t) = \textstyle\int_{-\infty}^{\infty} H(f) \cdot \mathrm{e}^{j2\pi f t} dt \\ & H(f) = \begin{cases} 1, & |f| \le \beta_r/2 \\ 0, & \text{other} \end{cases} \end{aligned} \tag{9.8}$$

where $S_{rs}(t)$ is:

$$\begin{aligned} s_{rs}(t) \approx & \frac{\sqrt{P_r}}{2} \cdot \frac{4}{\pi} \cdot \left[s_{AD}(t) + s_{AP}(t)\right] \\ & + \frac{\sqrt{P_r}}{2}\left[s_{BD}(t) + s_{BP}(t)\right]\left[\frac{4}{\pi} e^{j2\pi \cdot 2 f_s t} + \frac{4}{3\pi} e^{-j2\pi \cdot 2 f_s t}\right] \end{aligned} \tag{9.9}$$

The receiver's local reference signal is $S_{AP}(t)$, coherent integration time is T_p, and T_p is a integer multiple of PN code period of lower sideband pilot component. Taking no account of the impact of signal's received power, Then cross correlation function is:

$$\begin{aligned} R(\varepsilon) &= \tfrac{1}{T_P}\int_0^{T_P} s_r(t)s_{AP}^*(t-\varepsilon)dt = h(\varepsilon)\otimes R_{rs}(\varepsilon) \\ R_{rs}(\varepsilon) &= \tfrac{1}{\pi}\left(R_{AD,AP}(\varepsilon)+R_{AP,AP}(\varepsilon)+R_{BD,AP}(\varepsilon)+R_{BP,AP}(\varepsilon)\right) \end{aligned} \tag{9.10}$$

Because only the correlation function value of $|\varepsilon| < T_c$ is important when receivers is tracking signal, the assumption that $R(\varepsilon) = 0, \quad |\varepsilon| > T_c$ has no effect on code tracking performance analysis. Considering that the PN code length of pilot component is a integer multiple of the PN code length of data component. At the same time, assume that data message bit maintains unchanged during a integration period. We have the derivation that

$$\begin{aligned} R_{AD,AP}(\varepsilon) &= \tfrac{2}{T_P}\int_0^{T_P} s_{AD}(t)s_{AP}^*(t-\varepsilon)dt = \begin{cases} r_{AD,AP}(0)\cdot\frac{|\varepsilon|}{T_c}, & -T_c \le \varepsilon < 0 \\ r_{AD,AP}(1)\cdot\frac{|\varepsilon|}{T_c}, & 0 \le \varepsilon \le T_c \end{cases} \\ r_{AD,AP}(k) &= \tfrac{1}{N_{AP}}\sum_{n=0}^{N_{AP}-1} C_{AD}(n)\cdot C_{AP}(n+k) \end{aligned} \tag{9.11}$$

And

$$\begin{aligned} R_{AP,AP}(\varepsilon) &= \tfrac{2}{T_P}\int_0^{T_P} s_{AP}(t)s_{AP}^*(t-\varepsilon)dt = r_{AP,AP}(0)\cdot\left(1-\tfrac{|\varepsilon|}{T_c}\right), \quad -T_c \le \varepsilon \le T_c \\ r_{AP,AP}(k) &= \tfrac{1}{N_{AP}}\sum_{n=0}^{N_{AP}-1} C_{AP}(n)\cdot C_{AP}(n+k) \end{aligned} \tag{9.12}$$

$$\begin{aligned} R_{BD,AP}(\varepsilon) &= \frac{2}{T_P}\int_0^{T_P} s_{BD}(t)\left(\mathrm{e}^{j2\pi 2f_s t}+\frac{1}{3}\mathrm{e}^{-j2\pi 2f_s t}\right)s_{AP}^*(t-\varepsilon)dt \\ &= \begin{cases} r_{BD,AP}(0)\cdot\frac{\sin(2\pi f_s|\varepsilon|)}{2\pi f_s T_c}\left(\mathrm{e}^{-j2\pi f_s|\varepsilon|}+\frac{1}{3}\mathrm{e}^{j2\pi f_s|\varepsilon|}\right), & -T_c \le \varepsilon < 0 \\ r_{BD,AP}(1)\cdot\frac{\sin(2\pi f_s|\varepsilon|)}{2\pi f_s T_c}\left(\mathrm{e}^{j2\pi f_s(2T_c+|\varepsilon|)}+\frac{1}{3}\mathrm{e}^{-j2\pi f_s(2T_c+|\varepsilon|)}\right), & 0 \le \varepsilon \le T_c \end{cases} \\ r_{BD,AP}(k) &= \frac{1}{N_{AP}}\sum_{n=0}^{N_{AP}-1} C_{BD}(n)\cdot C_{AP}(n+k) \end{aligned} \tag{9.13}$$

$$\begin{aligned} R_{BP,AP}(\varepsilon) &= \frac{2}{T_P}\int_0^{T_P} s_{BP}(t)\left(\mathrm{e}^{j2\pi 2f_s t}+\frac{1}{3}\mathrm{e}^{-j2\pi 2f_s t}\right)s_{AP}^*(t-\varepsilon)dt \\ &= \begin{cases} \left(\begin{array}{l} r_{BP,AP}(0)\cdot\frac{\sin(2\pi f_s(T_c-|\varepsilon|))}{2\pi f_s T_c}\cdot \\ \left(\mathrm{e}^{j2\pi f_s(T_c-|\varepsilon|)}+\frac{1}{3}\mathrm{e}^{-j2\pi f_s(T_c-|\varepsilon|)}\right) \end{array}\right), & -T_c \le \varepsilon \le 0 \\ \left(\begin{array}{l} r_{BP,AP}(0)\cdot\frac{\sin(2\pi f_s(T_c-|\varepsilon|))}{2\pi f_s T_c}\cdot \\ \left(\mathrm{e}^{j2\pi f_s(T_c+|\varepsilon|)}+\frac{1}{3}\mathrm{e}^{-j2\pi f_s(T_c+|\varepsilon|)}\right) \end{array}\right), & 0 \le \varepsilon \le T_c \end{cases} \end{aligned} \tag{9.14}$$

From above we can see that $R_{AP,AP}(\varepsilon)$ is a real even function. The real part of $R_{BP,AP}(\varepsilon)$ is a even function, and its imaginary part is an odd function (because $2f_sT_c$ is a positive integer). Thus, whether C_{BP} is equal to C_{AP} or not doesn't affect the symmetry of $R_{rs}(\varepsilon)$. When $r_{AD,AP}(0)$ is equal to $r_{AD,AP}(1)$, $R_{AB,AP}(\varepsilon)$ is a real even function. When $r_{BD,AP}(0)$ is equal to $r_{BD,AP}(1)$, the real part of $R_{BD,AP}(\varepsilon)$ is a even function, and its imaginary part is an odd function. When these two conditions are met, the real part of $R_{rs}(\varepsilon)$ is a even function, and its imaginary part is an odd function. That is $R_{rs}(\varepsilon)$ is a symmetrical correlation function.

Figure 9.1 shows an example of asymmetrical correlation function. Let $r_{AD,AP}(0) = -1/1{,}023$, $r_{AD,AP}(1) = 63/1{,}023$, $r_{AP,AP}(0) = 1$, $r_{BD,AP}(0) = 63/1{,}023$, $r_{BD,AP}(1) = -1/1023$, $r_{BP,AP}(0) = 1$, that is C_{BP} is the same with C_{AP}. The received bandwidth is 20.46 MHz.

Figure 9.2 shows an example of symmetrical correlation function. Let $r_{AD,AP}(0) = 63/1{,}023$, $r_{AD,AP}(1) = 63/1{,}023$, $r_{AP,AP}(0) = 1$, $r_{BD,AP}(0) = -1/1{,}023$, $r_{BD,AP}(1) = -1/1{,}023$, $r_{BP,AP}(0) = 1$, that is C_{BP} is the same with C_{AP}. The received bandwidth is 20.46 MHz.

From the above analysis we can see that the selection of PN code for data components at lower and upper sideband has a significant impact on the symmetry of correlation function. Incorrect PN code correlation may result in an asymmetrical correlation function (as illustrated in Fig. 9.1), which would cause the cross zero point of discriminator function to move, so code tracking bias error is introduced. The selection of PN code for pilot components at upper sideband may change the shape of correlation function, and then the zero-crossing point slope of discrimination curve would change. So the selection of PN code for pilot components at upper sideband may affect code tracking performance and anti-multipath performance.

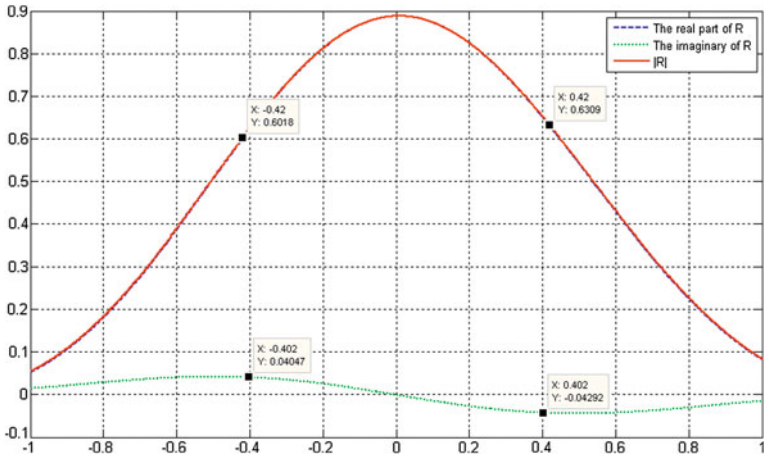


Fig. 9.1 The asymmetrical correlation function

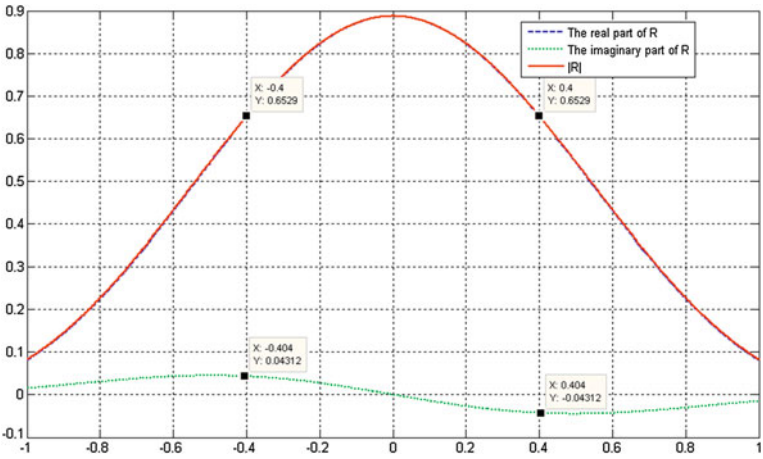


Fig. 9.2 The symmetrical correlation function

9.4 Simulation Results

First simulation results of zero-crossing point offset of discrimination curve are given out. Let GPS C/A code be the PN code of TD-AltBOC(15,10) signal. The PN code of data component at lower sideband is PRN 1 C/A code. The PN code of pilot component at lower sideband is PRN 2 C/A code. The PN code of data component at upper sideband is PRN 3 C/A code. The PN code of pilot component at upper sideband is PRN 2 or PRN 4 C/A code. The coherent integration time is 0.4 ms. Data message bit maintains unchanged during a integration period. The received bandwidth is 20.46 MHz. The sample rate is 600 MHz. As a comparison, Galileo AltBOC(15,10) signal uses the same parameters for simulation. The simulation results are shown in Fig. 9.3.

In Fig. 9.3a, $r_{AD,AP}(0) = -1/1{,}023$, $r_{AD,AP}(1) = -1/1{,}023$, $r_{AP,AP}(0) = 1$, $r_{BD,AP}(0) = -1/1{,}023$, $r_{BD,AP}(1) = -1/1{,}023$, $r_{BP,AP}(0) = -1/1{,}023$. When C_{BP} is the same with $C_{AP.}$, $r_{BP,AP}(0) = 1$.

In Fig. 9.3b, the PN code of data component at lower sideband is a code which is obtained by PRN 1 C/A code shifting left 52 chips. The PN code of pilot component at lower sideband is PRN 2 C/A code. The PN code of data component at upper sideband is a code which is obtained by PRN 3 C/A code shifting left 140 chips. The PN code of pilot component at upper sideband is PRN 2 or a code which is obtained by PRN 4 C/A code shifting left 15 chips. So $r_{AD,AP}(0) = 63/1{,}023$, $r_{AD,AP}(1) = 63/1{,}023$, $r_{AP,AP}(0) = 1$, $r_{BD,AP}(0) = -65/1{,}023$, $r_{BD,AP}(1) = -65/1{,}023$, $r_{BP,AP}(0) = 63/1{,}023$. When C_{BP} is the same with $C_{AP.}$, $r_{BP,AP}(0) = 1$.

In Fig. 9.3c, the PN code of data component at lower sideband is a code which is obtained by PRN 1 C/A code shifting left 10 chips. The PN code of pilot component at lower sideband is PRN 2 C/A code. The PN code of data component at upper sideband is a code which is obtained by PRN 3 C/A code shifting left 5 chips.

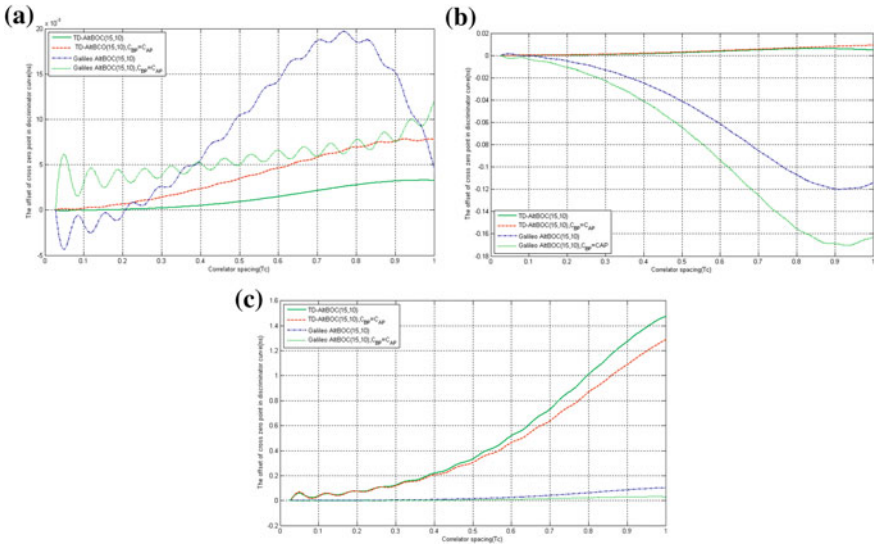


Fig. 9.3 The offset of cross zero point in discriminator curve. **a** symmetry case one, **b** symmetry case two, **c** asymmetry case

The PN code of pilot component at upper sideband is PRN 2 or a code which is obtained by PRN 4 C/A code shifting left 15 chips. So $r_{AD,AP}(0) = -1/1{,}023$, $r_{AD,AP}(1) = 63/1{,}023$, $r_{AP,AP}(0) = 1$, $r_{BD,AP}(0) = 63/1{,}023$, $r_{BD,AP}(1) = -1/1{,}023$, $r_{BP,AP}(0) = 63/1{,}023$. When C_{BP} is the same with $C_{AP.}$, $r_{BP,AP}(0) = 1$.

Figure 9.3a and b meet the symmetry condition in Sect. 9.3, so zero-crossing point offset of discrimination curve for TD-AltBOC(15,10) is less than 0.001 ns, which is close to zero and better than Galileo AltBOC(15,10) signal's. However, Fig. 9.3c doesn't meet the symmetry condition in Sect. 9.3, so zero-crossing point offset of discrimination curve for TD-AltBOC(15,10) is much larger than Galileo AltBOC(15,10) signal's, which is close to 1.4 ns. So the theory analysis results in Sect. 9.3 are verified.

Then we give out the simulation results of pilot component at lower sideband code tracking performance in thermal noise environment. The PN code is the same with those in Fig. 9.3b. The received bandwidth is 20.46 MHz. Single side code loop bandwidth is 1 Hz. The coherent integration time is 1 ms. The correlator space is 0.5 chips. The discriminator function is noncoherent Early minus late Power discriminator function. We consider two cases, PN codes of pilot component at lower and upper sideband are the same or not. The results are shown in Fig. 9.4. When these two pilot components' PN code are the same, the shape of correlation function changes a little. But the impact on code tracking performance is negligible.

Finally anti-multipath performance of pilot component at lower sideband is given out. Assume multipath direct ratio (MDR) is −6 dB. The correlator space is

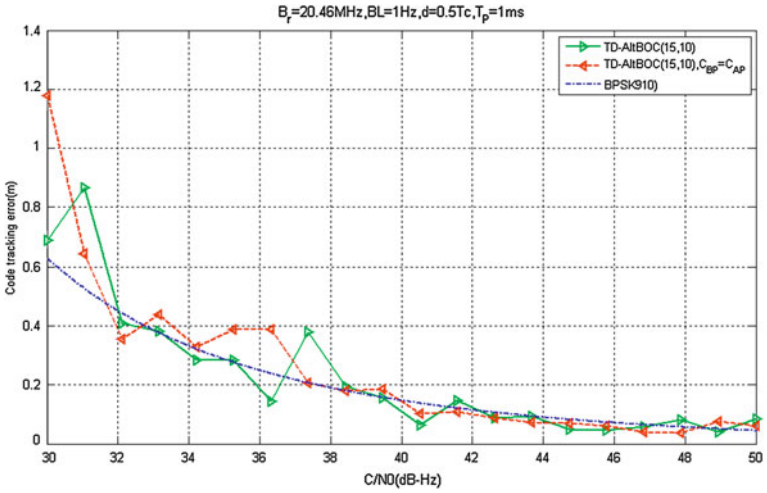


Fig. 9.4 Code tracking error versus C/N0

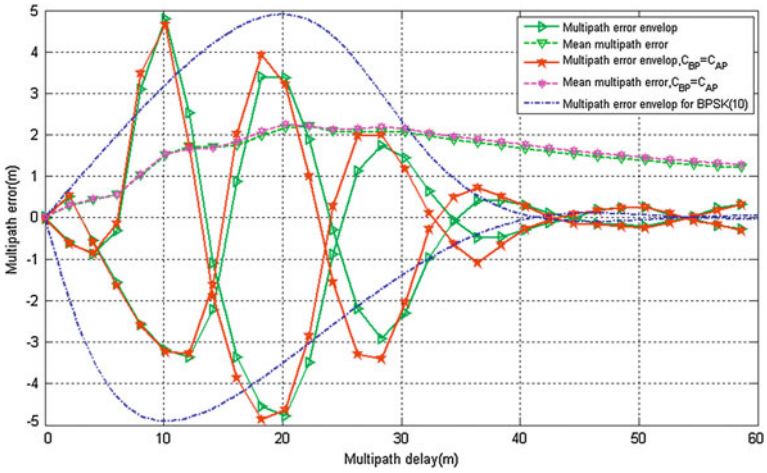


Fig. 9.5 Multipath error

0.5 chips. From Fig. 9.5 we can know that whether PN codes of pilot component at lower and upper sideband are the same or not, anti-multipath performance just changes a little.

9.5 Conclusions

In this paper, we make performance analysis on single sideband of TD-AltBOC modulation Signal. Specially, When the receiver is only tracking lower sideband signal's pilot component, due to the non-ideal PN codes and the effect of upper sideband signals, a asymmetrical correlation function may be obtained, which causes the cross zero point of discriminator function to move, so code tracking bias error is introduced. To solve this problem, we analyze the reason and condition why correlation function is asymmetry by theory analysis and simulation. Results show that when only considering TD-AltBOC signal from single satellite, the incorrect selection of data components' PN code at lower and upper sideband is the main reason leading to an asymmetrical correlation function. To ensure correlation function symmetry, the correlation values between the PN code of pilot component at lower sideband and the PN code of data component have to meet the following conditions: these two correlation values are equal when code phase delay is 0 chip and 1 chip. When these two pilot components' PN code are the same, the shape of correlation function changes a little. But the impact on code tracking performance is negligible. The main contribution of this paper is that our results can provide some useful guidance for the choose of PN codes in TD-AltBOC signal, which can make TD-AltBOC(15,10) signal's single sideband performance with Galileo AltBOC(15,10) signal similar. The results in this paper also provide references for other time-division GNSS signal.

References

1. European Union (2010) European GNSS (Galileo) open service signal in space interface control document. (OS SIS ICD, Issue 1.1, pp 4–7. http://ec.europa.eu/enterprise/policies/satnav/galileo/open-service/index_en.htm
2. Shivaramaiah NC, Dempster AG (2009). The Galileo E5 AltBOC: understanding the signal structure. In: International global navigation satellite systems society IGNSS symposium
3. Lestarquit L, Artaud G, Issler JL (2008) AltBOC for dummies or everything you always wanted to know about AltBOC. In: Proceedings of the 21st international technical meeting of the satellite division of the Institute of Navigation, pp 961–970
4. Tang ZP, Zhou HW, Hu XL, Ran YH, Liu YQ, Zhou YL (2010) Research on performance evaluation of compass signal. Scientia Sinica Phys Mech Astron 40(5):592–602
5. Tang ZP, Zhou HW, Wei JL (2011) TD-AltBOC: a new COMPASS B2 modulation. Sci China-Phys Mech Astron 54(6):P1014–P1021

Chapter 10
Research on GNSS Interoperability Parameters

Xiaochun Lu, Tao Han, Xue Wang and Fang Cheng

Abstract Interoperability of Global Navigation Satellite System (GNSS) is a hot topic in GNSS area, and a lot of excellent related researches have come into being. But, existence of differences among GNSSs cause inconvenience to users. In order to achieve interoperability and convenient for users, interoperability parameters are required. First of all, differences among GNSSs are defined based on their effects on users in position, velocity and timing (PVT). These differences may from message or/and signal, and are different in precision or/and format, which bring troubles when user is using the navigation resources from different GNSSs. To ameliorate the troubles caused by GNSS differences during PVT progress, the parameters which can represent these differences should be determined. These are ten parameters in this paper which form the GNSS interoperability parameter set, and each element of this set is defined as interoperability parameter. The interoperability parameters can be separated into signal level and message level, and should be broadcast to users in a common time reference and coordinate reference frame. Then, we discussed the definition, detection and calculation method of each interoperability parameter. At last, when broadcasting these parameters to users, there three different ways such as internet, mobile communication and commercial satellite, except the providers own links.

Keywords Interoperability · Interoperability parameters · Monitoring · Third party

X. Lu · T. Han (✉) · X. Wang · F. Cheng
National Time Service Center, Chinese Academy of Sciences, Xi'an, People's Republic of China
e-mail: hantao@ntsc.ac.cn

X. Lu · T. Han · X. Wang · F. Cheng
Key Laboratory of Precision Navigation and Timing Technology, Xi'an, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_10,

10.1 Introduction

Although interoperability incurs much focus among the Global Navigation Satellite System (GNSS) world, it still be an investigation state [1]. There is not a quite viable method to realize GNSS interoperability. The very reason that troubles the GNSS interoperability realization is the existent differences among each GNSSs. These differences make the user utilize satellite navigation resources difficultly. For instance, each system has their own coordinate system; in one hand, their biases will limit the positioning accuracy improvement of interoperability, in another hand, leave the coordinate transform process to users will greatly increase the complexity at users' level [2, 3].

Thus, there in dire need of the research on how to realize the interoperability exactly follow the definition of GNSS interoperability, which means enhance the service performances while do not increase the users' costs distinctly.

The key point of realize GNSS interoperability is to remove the differences among GNSSs, and to achieve this aim, we need:

- Define the differences among GNSSs;
- Study the parameters to represent these differences;
- Process and transmit the parameters to users;
- Make sure the users can depend on the parameters to improve services.

In this article, differences among GNSSs are defined based on their effects on users in position, velocity and timing (PVT), firstly. These differences may from message or/and signal, and are different in precision or/and format, which bring troubles when user is using the navigation resources from different GNSSs. To ameliorate the troubles caused by GNSS differences during PVT progress, the parameters which can represent these differences should be determined. These are ten parameters in this paper which form the GNSS interoperability parameter set, and each element of this set is defined as interoperability parameter. The interoperability parameters can be separated into signal level and message level, and should be broadcast to users in a common time reference and coordinate reference frame. Then, we discussed the definition, detection and calculation method of each interoperability parameter. At last, when broadcasting these parameters to users, there three different ways such as internet, mobile communication and commercial satellite, except the providers own links.

10.2 Differences Among GNSSs

10.2.1 Existent Differences

The existent GNSS differences [4] are represent in the flowing five aspects:

- Constellation
  - Satellite number, Types of satellite orbit, etc.
- Signal
  - Modulation, Center frequency, Received power, etc.
- Message
  - Massage structure, Data content, Data format, etc.
- System time reference
- System coordinates reference

10.2.2 Influence from GNSS Differences to Users

The influence from GNSS differences to users should been discovered in the positioning equation:

$$\begin{aligned}\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}-c\cdot\upsilon_{T_b}\\ =\tilde{\rho}_i+(\delta\rho_i)_{ion}+(\delta\rho_i)_{trop}-c\cdot\upsilon t_i^a\end{aligned} \tag{10.1}$$

where (x,y,z) is the unknown user location coordinates; (x_i, y_i, z_i) is the position coordinates of the ith satellite which is employed to solve the user location; c is the speed of light; υ_{T_b} is the correction data of receiver clock; υt_i^a is the correction data of the ith satellite clock; $\tilde{\rho}_i$ is the pseudo-range from the ith satellite to user; $(\delta\rho_i)_{ion}$ is the ionosphere time delay correction data of the ith satellite; $(\delta\rho_i)_{trop}$ is the troposphere time delay correction data of the ith satellite.

The elements (x_i, y_i, z_i), $\tilde{\rho}_i$, $(\delta\rho_i)_{ion}$ and υt_i^a are different because the GNSS differences. These elements can be obtained from signals or messages of different GNSSs; and have different format and/or precision. We should determine the type of these differences firstly, see Table 10.1.

Table 10.1 Source and representation of different elements

Elements	Differences			
	From signal	From message	Format	Precision
(x_i, y_i, z_i)		√	√	√
$(\delta\rho_i)_{ion}$	√	√	√	√
$\tilde{\rho}_i$	√	√	√	√
υt_i^a		√		√

10.3 Function of GNSS Interoperability Parameter

10.3.1 Interoperability Parameter Set

The elements in Table 10.1 are determined by many factors, see Table 10.2.

Those items in Table 10.2 form the GNSS interoperability parameter set, and each of the items is the GNSS interoperability parameters.

10.3.2 Realize Ideal GNSS Interoperability via GNSS Interoperability Parameters

GNSS differences are the very reason limits the service performance improvement, and the essential reason that increases the complexity of using satellite navigation resources from different GNSSs [5, 6].

In Fig. 10.1, when a user prepares to use the multi-GNSS resources to realize PVT, there are too many data from each GNSSs that represent the same physical meaning.

Thus, there requires a method to transform the different parameters into the same type, which means the same format and same precision level, see Fig. 10.2.

By using GNSS interoperability parameters, there is no more distinguish among GNSSs. In the user's view, the satellites are the same, with the common GNSS satellites, common GNSS signal, common GNSS ionosphere, common GNSS time and common GNSS coordinate (Fig. 10.3).

10.3.3 How to use GNSS Interoperability Parameters

In order to provide better service for users, we needs:

- Utilize specific method to monitor interoperable parameters;

Table 10.2 GNSS differences and their items

Differences	Items
Signal	User-received signal level, modulation error, correlation characteristics, phase coherence, TGD
Ephemeris	Orbit offset, GNSS reference bias
Onboard Clock	Clock offset, GNSS time bias
Propagation	Ionosphere
……	……

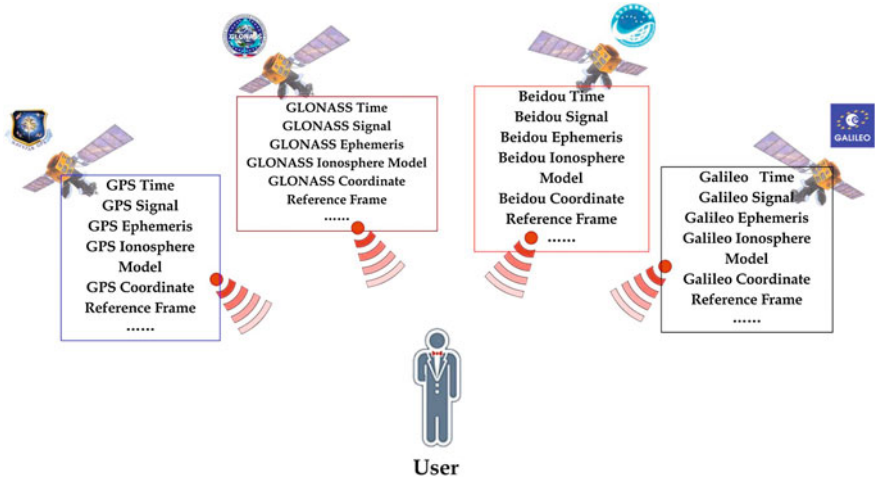


Fig. 10.1 GNSS interoperability without interoperability parameters

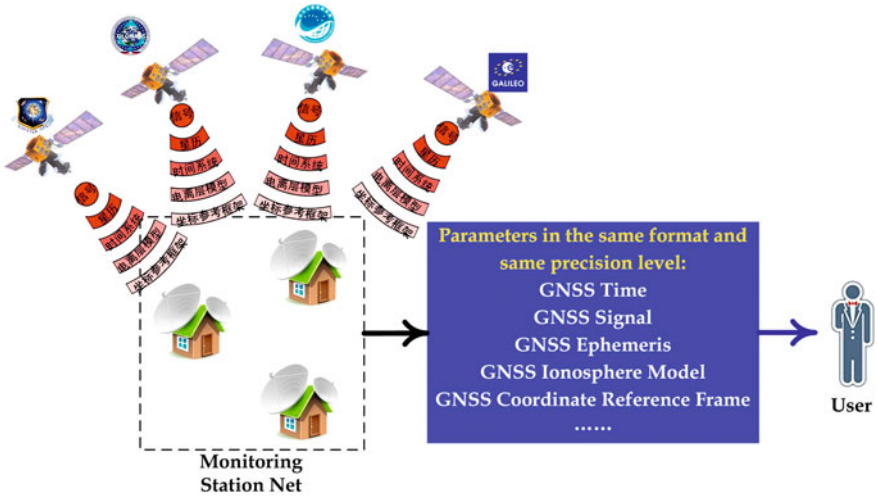


Fig. 10.2 Transform different parameters into the same type

- Calculate parameters in a common time reference and coordinate reference frame;
- Broadcast parameters to users;
- Take parameters to eliminate adverse effect of the GNSS differences (Fig. 10.4).

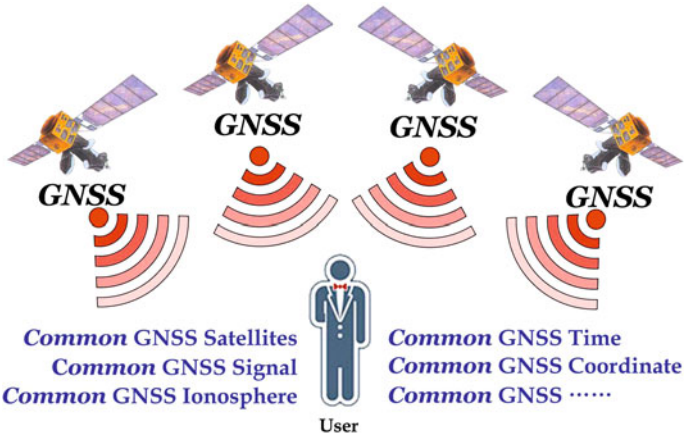


Fig. 10.3 Ideal GNSS interoperability

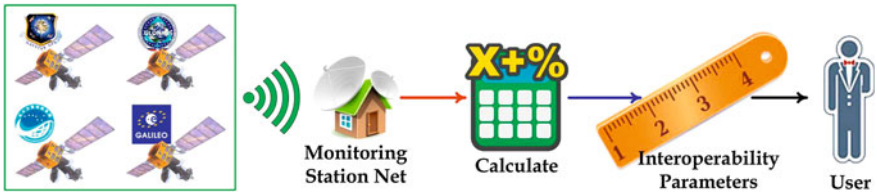


Fig. 10.4 Flow chart of using interoperability parameters

10.4 GNSS Interoperability Parameters

According to Table 10.1, there are ten parameters in the GNSS interoperability set: User-Received Signal Level, Modulation Error, Correlation Characteristics, Phase Coherence, TGD, Orbit offset, GNSS reference bias, Clock offset, GNSS Time Bias and Ionosphere. These parameters can be divided into two sub-sets, see Fig. 10.5.

In the following part, the definition, detection method and calculation method of each parameter will be discussed.

10.4.1 Signal Parameters

Provide users interoperable parameters in signal level will let the users have the ability to chose a high quality signal in receiving process, which will reduce the first positioning duration and decrease the complexity of receiver.

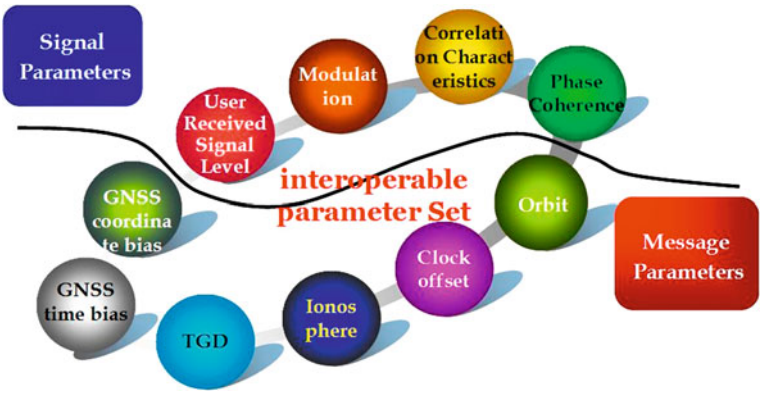


Fig. 10.5 Two sub-sets of GNSS interoperability parameter sets

10.4.1.1 User-Received Signal Level

The definition of User-Received Signal Level is: the signal power when it arrives at the ground station.

The detection method: use ground monitoring receiver to monitor the signal power and its variation range.

10.4.1.2 Modulation Error

The definition of Modulation Error: the error which was produced in the modulation and transmission, including: phase modulation error and amplitude modulation error.

Phase modulation error: the differences between the real phase in each channel and the ideal phase of the signal.

Amplitude modulation error: the differences between the real amplitude in each branch and the ideal amplitude of the signal.

The detection method: compare the received signal to designed signal.

10.4.1.3 Correlation Characteristics

The definition of Correlation Characteristics: outputs of correlation peak amplitude and correlation curve characteristics after the operation of signal correlation.

Correlation loss: power difference between the actually received signal and the ideal signal in the designed bandwidth of the signal.

Correlation curve: the curve obtained through correlation calculation between recovered ranging code and the ideal ranging code of all signals.

The detection method: the monitoring receiver acquire navigation signal, and then evaluate amplitude attenuation and curve distortion which is caused by wave distortion.

Calculation method: the correlation function is shown as Eq. (10.2):

$$CCF(\varepsilon) = \frac{\int_0^{T_p} S_{BB} - PreProc(t) \cdot S_{Ref}(t-\varepsilon)dt}{\sqrt{\left(\int_0^{T_p} |S_{BB} - PreProc|^2 dt\right) \cdot \left(\int_0^{T_p} |S_{Ref}(t)|^2 dt\right)}} \tag{10.2}$$

$S_{BB} - PreProc$ is the base-band signal been pretreated (down conversion, Doppler removal); reference signal S_{Ref} is ideal base-band signal generated by local receiver; integral time T_p is the main code period of reference signal.

Relative loss, power loss of available signal to all received signals:

$$P_{CCF}[dB] = \max_{over\ all\ \varepsilon} (20 \cdot \log_{10}(|CCF(\varepsilon)|)) \tag{10.3}$$

10.4.1.4 Phase Coherence

The definition of Phase Coherence: the relative change of signal elements in the timeline. There are code-carrier coherence and codes coherence.

Code-carrier Coherence: Relative jitter value between ranging code and carrier wave in the same signal branch.

Codes Coherence: Relative jitter value of time delay between ranging code and carrier wave; relative jitter value of ranging codes in different signal branch.

The detection method is to monitor the navigation signal.

Firstly, the code-carrier coherence calculation is shown as Eq. (10.4).

$$\begin{aligned} CCD^{L_i}_{L_j,L_k}(t,\, t+T) = {} & PR_{Li}(t+T) - PR_{Li}(t) \\ & - [CR_{Li}(t+T) - CR_{Li}(t)] \\ & - 2\left(\frac{f_{L_1}}{f_{L_i}}\right)^2 \Delta I_{Lj,L_k}(t,\, t+T) \end{aligned} \tag{10.4}$$

$\Delta I_{Lj,L_k}(t,\, t+T)$ denotes the ionospheric delay differences in L1 frequency on interval $[t,\, t+T]$, these differences calculated from D-value of L_j and L_k frequency amplitude. If $CCD^{L_i}_{L_j,L_k}(t,\, t+T)$ meets:

$$100 \le T \le 7,200,\ t_1 \le t \le t_2 - T,\ CCD^{L_i}_{L_j,L_k}(t, t+T) > 6.1\ \text{m} \tag{10.5}$$

Thus, code and carrier wave are consistence at $t+T$.

Then we will discuss the calculation method of coherence in ranging codes.

$$\begin{aligned}\rho_i' &= \rho_i + \lambda_i^2 \frac{\Phi_j - \Phi_i}{\lambda_j^2 - \lambda_i^2},\\ \rho_j' &= \rho_j + \lambda_j^2 \frac{\Phi_i - \Phi_j}{\lambda_i^2 - \lambda_j^2},\\ \Delta\rho &= \rho_i - \rho_j - \Delta\rho_{i,j.}\end{aligned} \tag{10.6}$$

$\rho_i', \rho_j', \rho_i, \rho_j$ represent pseudo-range when exist ionospheric error and no ionospheric error respectively, j represent frequency which is different from i. Φ_i and Φ_j are observation value of carrier phase (the unit is distance), wavelength are λ_i and λ_j respectively; $\lambda_i^2(\Phi_j - \Phi_i)/\lambda_j^2 - \lambda_i^2$ is amended value of dual-frequency ionosphere; $\Delta\rho$ represent time delay between receiver channels.

10.4.2 Message Parameters

Function of information level parameters is to eliminate adverse effect of the GNSS differences, and provide convenient service for users, see Fig. 10.6.

10.4.2.1 Orbit Offset

Definition of Orbit offset: GNSS precise orbit calculated based on same monitor station, same orbit determination algorithm, and same space–time reference.

Detection method: utilize observation value of multi-mode receiver and precise orbit algorithm to calculate GNSS precise orbit.

Calculation process of Orbit offset:

- Detect the coarse error of observation value and cycle slip;
- Using the processed data to the precision of satellite orbit, station location and ERP parameter estimation;
- Obtained by compare to the correction information of broadcast ephemeris orbit.

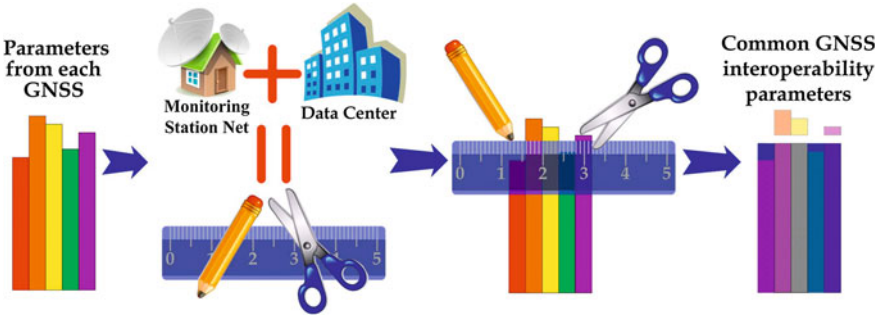


Fig. 10.6 Function of message parameters

10.4.2.2 Clock Offset

Definition of Clock offset: calculate GNSS precise clock error based on same monitor station, same orbit determination algorithm, and same space–time reference.

Detection method: utilize Observation data of Laser, radio, dual-frequency carrier wave and precise clock error algorithm to get precise clock error of GNSS.

Calculation process of Clock offset:

- Detect the coarse difference and cycle slip of the observation data;
- Take real-time precise satellite orbit, the position of observation station, and EPR parameters as known parameters to real-time precise satellite clock error processor;
- Use preprocessed real-time observation data by means of Square Root Filter to evaluate clock error;
- Compare the evaluated clock error with the broadcasted clock error, and then get amended clock error information.

10.4.2.3 Ionosphere

Definition of Ionosphere: total electron content of global ionospheric grid based on the calculation of dual-frequency observations.

Detection method: by monitoring of ionospheric grid model.

Calculation process of Ionosphere:

- Total electron content (TEC) in the path from monitoring station to satellite:

$$\tilde{P}^k_{1,i} - \tilde{P}^k_{2,i} = (1-\xi)\frac{40.28 \cdot TEC}{f_1^2} + \Delta b^k + \Delta b_i \tag{10.7}$$

- Work out total electron content of global ionospheric grid by geomagnetic model:

$$TEC(\phi,\lambda) = \sum_{n=0}^{n_d\max}\sum_{m=0}^{n}\tilde{P}_{nm}(\sin\ \phi)\cdot(\tilde{A}_{nm}\cos(m\lambda) + \tilde{B}_{nm}\sin(m\lambda)). \tag{10.8}$$

10.4.2.4 GNSS Time Bias

Definition of GNSS Time Bias: the time differences between each satellite navigation system and UTC.

Detection method: monitor each system time and compare with UTC.

Calculation process:

- Broadcasting method of interoperability parameters

$$\Delta_{sys} = T_{sys} - UTC(k) \tag{10.9}$$

- Worked out the difference Δk between $UTC(k)$ and UTC,

$$\Delta k = UTC(k) - UTC \tag{10.10}$$

- Normalized each system time to UTC,

$$\Delta = \Delta_{sys} + \Delta k \tag{10.11}$$

10.4.2.5 GNSS Reference Bias

Definition of GNSS reference bias: differences between coordinate reference frame and ITRF.

Detection method: measure the coordinate of given points in different coordinate reference frame, then calculate their difference.

Calculation: denotes $(X_n, Y_n, Z_n)_i^T$ as the coordinate of P_n in frame i, $(X_n, Y_n, Z_n)_{ITRF}^T$ as the coordinate of P_n in ITRF; then (using Bursa model):

$$\begin{aligned}
\Delta P_{n,(\mathrm{ITRF},i)} &= \begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix}_{\mathrm{ITRF}} - \begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix}_i \\
&= \begin{bmatrix} dX_n \\ dY_n \\ dZ_n \end{bmatrix} + \begin{bmatrix} dm & \varepsilon_3 & -\varepsilon_2 \\ -\varepsilon_2 & dm & \varepsilon_1 \\ \varepsilon_2 & -\varepsilon_3 & dm \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_i \\
\Delta_{\mathrm{ITRF},i} &= E(\Delta P_{(\mathrm{ITRF},i)}) \pm k \times \sqrt{D\left(P_{(\mathrm{ITRF},i)}\right)^2},\ k = 1,2,3
\end{aligned} \tag{10.12}$$

10.4.2.6 TGD

TGD is the GNSS signal group time delay, which can be monitored and compared the signal time delay. It is calculated by the following equation:

$$TGD(f_i, f_j) = \frac{PR(f_i) - PR(f_j)}{1 - (f_i, f_j)^2} \tag{10.13}$$

where, f_i and f_j are carrier wave frequency of two GNSS signals, $PR(f_i)$ and $PR(f_j)$ are the corresponded signal group time delay.

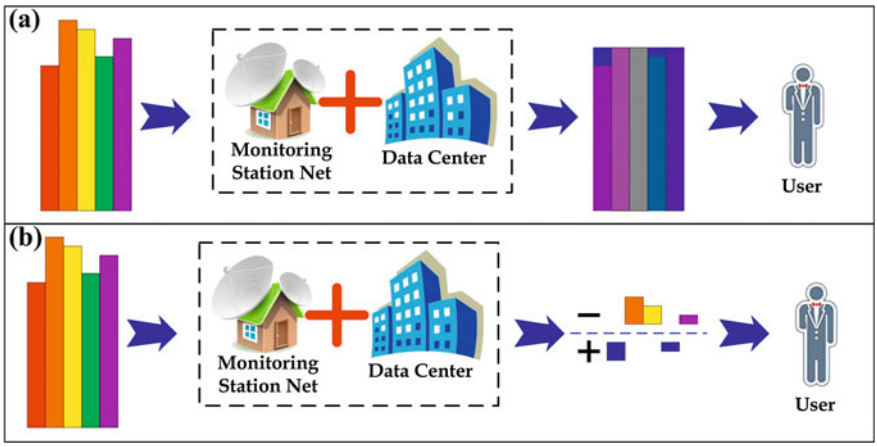


Fig. 10.7 Two broadcasting methods

10.5 Broadcasting Method of Interoperability Parameters

There are two method of interoperability parameter broadcasting (see Fig. 10.7):

- Broadcast parameters with uniform format and precision after the process of monitoring and calculating [7];
- Broadcast the differences between GNSS interoperability parameters and corresponded parameters in each system.

Each method has its own advantages and disadvantages, shown in Table 10.3.

Except the providers own links, interoperable parameters can be broadcasted to users in different ways: internet, mobile communication, commercial satellite (Fig. 10.8).

Table 10.3 Advantages and disadvantages of interoperability parameters broadcasting methods

	Advantages	Disadvantages
Precision data	1. Less computational complexity of user receiver 2. Less computational complexity of the third monitoring station	1. Need high data rate 2. Increased the time of receiving complete information 3. The system will be useless when the parameters of the third party are unavailable
Tolerance data	1. Low requirements of data rate 2. Receive complete information in short time 3. Can still use original GNSS to realize PVT when the parameters of the third party are unavailable	1. Additional process at user receiver 2. Increased the amount of computation in the third monitoring station

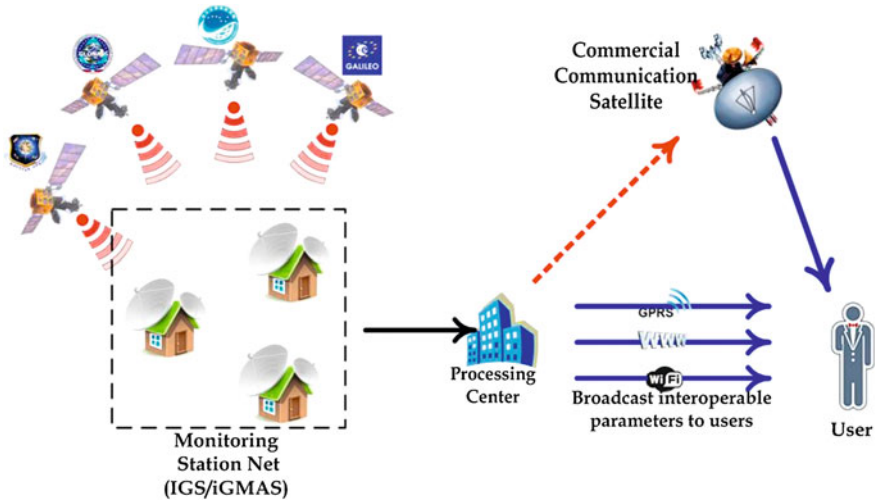


Fig. 10.8 GNSS interoperability realization

10.6 Conclusion

In this paper, we discussed the GNSS interoperability parameters, which can be employed to realize GNSS interoperability. There are still a lot of works to do in GNSS interoperability parameters, such as how to use these parameters in the GNSS receiver, and so on. The interoperability parameter will also lead a new research aspect of GNSS.

References

1. Report of the first meeting of the international committee on global navigation satellite systems (ICG) Providers forum. In: The first meeting of ICG, Bangalore, 4 Sept 2007
2. Stupak G (2008) The Russian federation view on GNSS compatibility and interoperability. In: The third meeting of the international committee on global navigation satellite systems (ICG), Pasadena, USA, 8–12 Dec
3. Han T, Lu X et al (2011) Differential equation dynamical system based assessment model in GNSS interoperability. Sci China Phys Mech Astron 54(6):996–1003
4. Lu J, Yang Q (2010) Study on GNSS interoperability. Sci China Phys Mechan Astron 40(5):1–8
5. Li J, Li Z, Hao J et al (2009) A preliminary study on compatibility and interoperability of GNSS. J Geometrics Sci Technol 26(3):177–180
6. Han T, Lu X, Rao Y et al (2011) Saturated GDOP value analysis in satellite navigation constellation. CSNC 2011, Shanghai
7. Lu X, Lu J, Bai Y et al (2010) Interoperability feasibility analysis between Beidou and GPS. Global navigation satellite systems: report of a joint workshop of the National Academy of Engineering and the Chinese Academy of Engineering, 2011, pp 75–82

Chapter 11
Impact Analysis of Navigation Signal in the Radio Astronomy Band

Hui-chao Zhou, Peng Li, Jin-jun Zheng, Zhong-gui Chen and Jian Wang

Abstract The band 1610.6–1613.8 MHz is used by the radio astronomy (RA) community in order to receive, measure and study the radio from star for the formation of protostars and the initial stages of star formation. The observation of the radio signal is very sensitive to the interference environment present in the band since the investigated signal is received with a very low power. In this paper, according to the International Telecommunications Union recommendations and characteristics of navigation system signal and constellation, a simulation is made which calculates the power rejection of the navigation signal in RA band. The result of this paper can provide a reference for the design of payload filter and the analysis of system compatibility.

Keywords Navigation satellite · Radio astronomy · Power rejection

11.1 Introduction [1]

The spectral line of the Hydroxyl (OH) radical transitions are considered to be one of the important references for the observation of physical phenomena associated with the formation of protostars and the initial stages of star formation, so the band 1610.6–1613.8 MHz is used by RA for research. The observation of the OH radical transition realized around 1612 MHz is very sensitive to the interference, and the L-band GNSS signals are almost all around 1612 MHz, so the interference from navigation signals cannot be ignored.

H. Zhou (✉) · P. Li · J. Zheng · Z. Chen · J. Wang
Institute of Space System Engineering, China Academy of Space Technology, Beijing 100094, People's Republic of China
e-mail: saltegg1986@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_11,

RA is a primary user of the band and thus, to protect the RA observations, the ITU has made several recommendations, including some dedicated to RNSS systems:

- For Geo-Stationary Orbit (GEO) satellites, the resolution ITU-R 739 considers the maximum Power-Flux Density (pfd) level to be -194 dBW/m^2/20 kHz for a single dish antenna.
- For Non-Geo-Stationary Orbit (non-GEO, especially MEO) satellites, the resolution ITU-R 739 considers the maximum Equivalent Power-Flux Density (epfd) level to be -258 dBW/m^2/20 kHz in 98 % time for a single dish antenna.

GPS L1 signal, GLONASS G1 signal and Galileo E1 signal provide interference in the RA band. Lots of research, evaluation and payload optimization work have been made in order to reducing or avoid the interference to the RA band. GNSS systems are being in the stage of building of updating, which make the number of satellite and signal band increasing quickly, so the signal design and interference evaluation is very important while the frequency resource is limited.

Based on the ITU recommendations to RNSS, the main goal of this paper is making an epfd calculation of single satellite and navigation constellation in order to achieve the simulation and evaluation of modern navigation signal in the RA band. The result of evaluation, which provides a reference value of power rejection of navigation signal in the RA band, is a combination simulation of the design of constellation, signal and payload. Because the realization of signal and constellation are based on the serving modern GNSS systems, the simulation process and result can be used as a kind of reference for payload design and system compatibility.

11.2 Analysis of GNSS Signal in the RA Band

Modern GNSS signal is achieved by BOC modulation. The analysis model is based on the structure of Galileo E1 signal and GPS L1 signal, which have center frequency at 1,575.42 MHz, MBOC and BOC sub-carrier modulation and code rate of 1.023 Mcps. So the signal structure [2] in the analysis model is shown as Table 11.1.

As shown in Fig. 11.1, the center frequency of model signal is 1575.42 MHz which has only 35.18 MHz interval to the RA band. In order to meet the ITU RA

Table 11.1 Structure of navigation signal

Signal	Center frequency (MHz)	Code rate (MHz)	Sub-carrier
S1_1	1,575.42	1.023	MBOC (6,1,1/11)
S1_2		1.023	
S2		2.046	BOC (14,2)

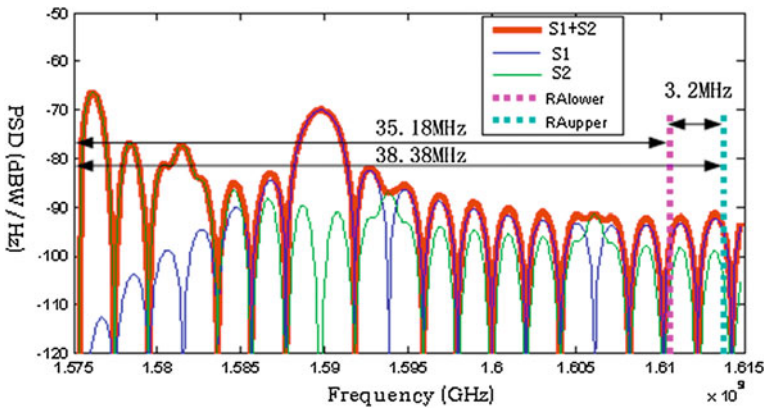


Fig. 11.1 Power spectral density of navigation signal. The signs of signals is the same in the Table 11.1

protection requirements and calculate the power rejection value of signal at the RA band, an analysis of the signal in the RA band is necessary.

1. Convert EIRP to pfd

The radius used by calculating pfd is the distance in the vector direction of the satellite and the center of the earth. The conversion equation of EIRP and pfd is shown as Eq. (11.1).

$$pfd = EIRP - [4\pi d^2] \text{ dB(W)} \tag{11.1}$$

In the process of analysis, both GEO satellite and MEO satellite should be considered. The EIRP of GEO satellite whose altitude is 36,000 km is 34 dBW which consults receiving power at ground and space attenuation. The pfd of GEO satellite can be calculated as the Eq. (11.1).

$$\begin{aligned} pfd_G &= EIRP_G - [4\pi d_G^2] \\ &= 34 - 162.12 = -128.12 \text{ dB}\left(\text{W/m}^2\right) \end{aligned} \tag{11.2}$$

The EIRP of MEO satellite whose altitude is 21,500 km is 32 dBW which consults receiving power at ground and space attenuation. The pfd of MEO satellite can be calculated as the Eq. (11.1).

$$\begin{aligned} pfd_M &= EIRP_M - [4\pi d_M^2] \\ &= 32 - 157.64 = -125.64 \text{ dB}\left(\text{W/m}^2\right) \end{aligned} \tag{11.3}$$

2. Convert the pfd in the whole signal band to the pfd in 20 kHz of the RA band

Because in the ITU recommendation, the requirement of pfd is the pfd in 20 kHz, not in the whole band, the convertion of the pfd in the whole signal band to the pfd

in 20 kHz of the RA band is necessary. As shown as Eq. (11.4), R(f) is the ratio of pfd in the every 20 kHz of the RA band to the pfd in the whole signal band. Parameter f is the distance to the center frequency of GNSS signal, which is in the range of 35.18–38.38 MHz.

$$R(f) = \int_{f}^{f+20\,\mathrm{kHz}} \frac{1}{2} GBOC_{(14,2)}(f)df + \int_{f}^{f+20\,\mathrm{kHz}} \left(\frac{1}{22} GBOC_{(6,1)}(f) + \frac{10}{22} GBOC_{(1,1)}(f) \right) df \tag{11.4}$$

GBOC(f) in the Eq. (11.4) is the normalization of pfd function. R(f) in the RA band is shown as Fig. 11.2. In order to ensure every 20 kHz pfd of the model signal in the RA band meet the ITU requirement, the maximum of R(f), shown as the mark R_{MAX} −48.34 dBi in Fig. 11.2, will be used in the analysis.

Considering the signal of single GEO satellite in the RA band, the maximum of pfd in 20 kHz is calculated in Eq. (11.5).

$$\begin{aligned} pfd_{G20\,\mathrm{kHz}} &= pfd_G + R_{\mathrm{MAX}} = -128.12 - 48.34 \\ &= -176.46 \ \ \mathrm{dB}\left(\mathrm{W/m^2/20\,kHz}\right) \end{aligned} \tag{11.5}$$

Considering the signal of single MEO satellite in the RA band, the maximum of pfd in 20 kHz is calculated in Eq. (11.6).

$$\begin{aligned} pfd_{M20\,\mathrm{kHz}} &= pfd_M + R_{\mathrm{MAX}} = -125.64 - 48.34 \\ &= -173.98\,\mathrm{dB}\left(\mathrm{W/m^2/20\,kHz}\right) \end{aligned} \tag{11.6}$$

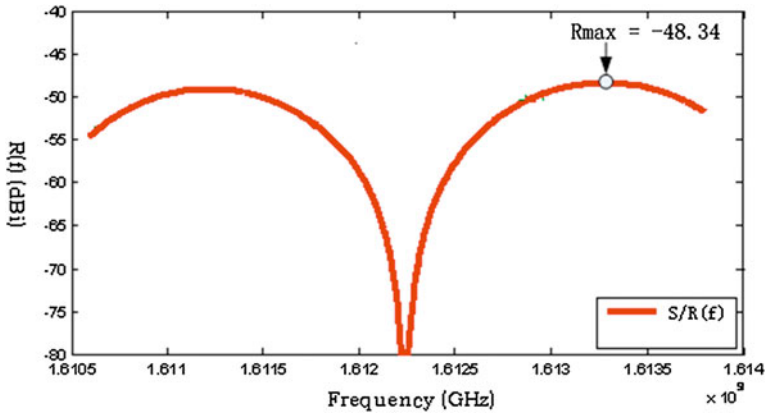


Fig. 11.2 Ratio of navigation signal power in 20 kHz of RA band and in the whole bandwidth

11.3 Discompose the epfd of Constellation into the pfd of Single Satellite

The position of GEO satellite will not change to the earth, so the interference from GEO satellite is constant. The ITU RA protection recommendation only requires the pfd of a single GEO is not exceed −194 dBW/m^2/20 kHz. So the design of GEO satellite payload should meet the pfd requirement in the RA band.

The position of MEO satellite is changing to the earth in a cyclical manner, what makes the interference to RA band from every single MEO satellite is also changing in the same period. The ITU RA protection recommendation requires the epfd of the whole constellation is not exceed −258 dBW/m^2/20 kHz in 98 % of the time. So in order to achieve the design threshold of a single MEO satellite payload, a whole period of MEO satellite constellation simulation is necessary [3].

According to the ITU-R M.1583-1, constellation epfd equation is as the Eq. (11.7) [4].

$$epfd = 10\log_{10}\left[\sum_{i=1}^{N_a} 10^{\frac{P_i}{10}} \cdot \frac{G_t(\theta_i)}{4\pi d_i^2} \cdot \frac{G_r(\varphi_i)}{G_{r,\max}}\right] \tag{11.7}$$

where:

Na number of non-GEO space stations that are visible from the radio telescope

i index of the non-GEO space station considered

Pi RF power of the unwanted emission at the input of the antenna (or RF radiated power in the case of an active antenna) of the transmitting space station considered in the non-GEO system (dBW) in the reference bandwidth

θi off-axis angle (degrees) between the boresight of the transmitting space station considered in the non-GEO system and the direction of the radio telescope

Gt(θi) transmit antenna gain (as a ratio) of the space station considered in the non-GEO system in the direction of the radio telescope

di distance (m) between the transmitting station consideredin the non-GEO system and the radio telescope

ϕi off-axis angle (degrees) between the pointing direction of the radio telescope and the direction of the transmitting space station considered in the non-GEO system

Gr(ϕi) receive antenna gain (as a ratio) of the radio telescope, in the direction of the transmitting space station considered in the non-GEO system (see Recommendation ITU-R RA.1631)

Gr,max maximum gain (as a ratio) of the radio telescope.

In the process of simulation, the gain pattern of sending antenna is a reference model. The receiving antenna refers to the 100 m diameter antenna located in

Germany (50.7°N, 7°E), whose gain pattern refers to ITU-R RA.1613 [5]. In this simulation, the bore sight of the receiving antenna points to the zenith and the minimum visible elevation of satellite is 0°. The gain pattern of RA 100 m antenna is shown in Fig. 11.3.

According to the simulation, the epfd samples of the receiving antenna in a constellation period can be calculated. The relationship between the pfd of single satellite and time percentage of epfd which can meet the requirement is shown in Fig. 11.4. The pfd threshold of single MEO satellite can be get from the marks in Fig. 11.4. So the pfd design of MEO satellite should not exceed −214 dBW/m^2/20 kHz in the RA band when the RA receiving antenna is the one in Germany.

The RA receiving antennas have been located in around the earth, and the GNSS signals also have been almost received around the whole world. So different places the RA antenna locate, different pfd results will be get. In this paper, an analysis, which places the RA antenna at every 5° of longitude and latitude on the whole north earth, is made to make sure the pfd result of MEO satellite is universal. The result on south earth is the same as north earth.

As shown in Fig. 11.5, in order not to disturb the 100 m RA antenna on earth observing the OH radical, the design pfd of single MEO satellite payload should not exceed −216 dBW/m^2/20 kHz in the RA band.

11.4 Power Rejection of the Signal in the RA Band

Based on the constellation and signal structure, the power rejection in the band 35.18–8.38 MHz off to center frequency of the signal should meet the condition as the Eq. (11.8).

$$pfd_{single-sat} + Att \leq pfd_{index} \tag{11.8}$$

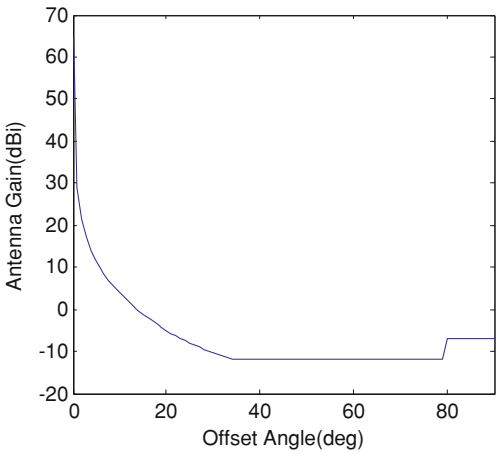


Fig. 11.3 RA 100 m antenna gain pattern

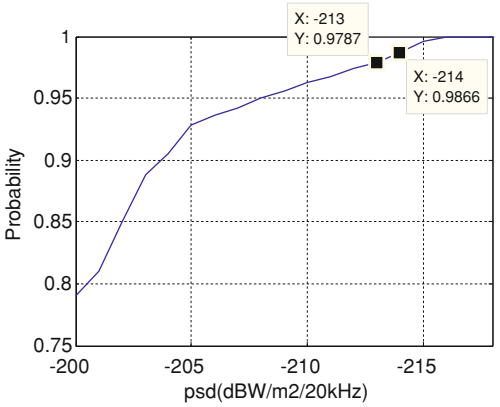


Fig. 11.4 PFD threshold of single satellite to RA antenna in Germany

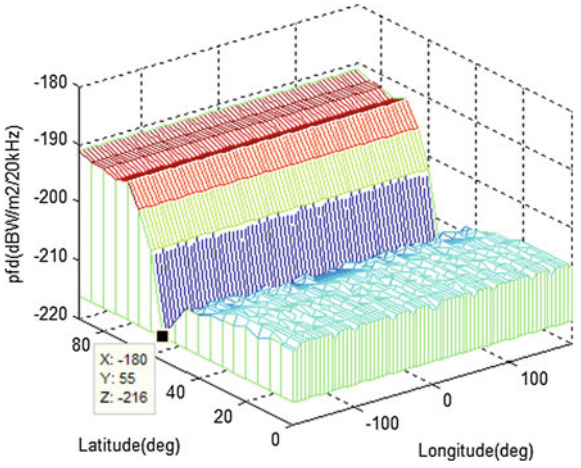


Fig. 11.5 PFD threshold of single satellite to RA antenna on north earth

where:

$pfd_{single-sat}$	The pfd (W/m^2/20 kHz) of the single satellite in the RA band
Att	The power attenuation of the signal needed to meet the ITU index in the RA band
pfd_{index}	The index in the ITU recommendation.

According to the results of Eqs. (11.5) and (11.6), in the model constellation, the maximum pfd of GEO satellite signal is −176.46 dBW/m^2/20 kHz and MEO satellite signal is −173.98 dBW/m^2/20 kHz.

In the ITU recommendation, the upper limit of single GEO satellite is −194 dBW/m^2/20 kHz and MEO satellite is −216 dBW/m^2/20 kHz. So the power rejection of GEO satellite in the analysis model can be calculated in Eq. (11.9).

Table 11.2 Parameters of navigation signal power rejection

Parameter	Satellites		
	GEO	MEO	Remarks
EIRP (dBW)	34	32	Radiate power
pfd (dBW/m^2)	−128.12	−125.64	Full span
Rmax (dBi)	−48.34		The pfd ratio of 20 kHz in the RA band to full band
pfd (dBW/m^2/20 kHz)	−176.46	−173.98	
epfd index (dBW/m^2/20 kHz)	≤ -194	≤ -258	Recommendations of ITU
pfd index of single satellite (dBW/m^2/20 kHz)	≤ -194	≤ -216	
Power rejection Att (dB)	≤ -17.54	≤ -42.02	In the RA band

$$\begin{aligned} Att_{\mathrm{G}} &\leq pfd_{Gindex} - pfd_{G20\,\mathrm{kHz}} \\ &= -194 - (-176.46) = -17.54\,\mathrm{dB} \end{aligned} \tag{11.9}$$

The power rejection of MEO satellite in the analysis model can be calculated in Eq. (11.10).

$$\begin{aligned} Att_{M} &\leq pfd_{Mindex} - pfd_{M20\,\mathrm{kHz}} \\ &= -216 - (-173.98) = -42.02\,\mathrm{dB} \end{aligned} \tag{11.10}$$

In the simulation, there is not a continuous equation about the gain pattern of the sending antenna, so it referred to some discrete sampling values which would make the result include about 1 dB error that is in the range of fluctuation of payload signal power.

11.5 Conclusion

The paper has done quantitative analysis to the interference of the navigation signal in the band 1610.6−1613.8 MHz. Then, we make some simulation to the equivalent power flux density of the signal of modern navigation system, according to the recommendations of ITU. With the requirement of ITU and the satellites constellation, the minimum power rejection of the GEO satellite signal in the RA band is 17.54 dB, and MEO satellite is 42.02 dB. The key parameter of the analysis is shown in Table 11.2.

According to the general characteristic of the signal structure of the modern GPS and Galileo system, the paper has done modeling analysis of the navigation signal and satellite constellation. The paper can provide a reference for the design of payload filter and the analysis of GNSS system compatibility.

References

1. Julien O, Issler J-L (2008) Mitigating the impact of GNSS signals in the radio astronomy band 1610.6–1613.8 MHz. ION GNSS 21st International technical meeting of the satellite division. pp 921–932
2. He C, Lu X, Wang X (2009) BOC modulation and simulation for GNSS. J Time Freq 32(2):134–141
3. Kaplan ED, Hegarty CJ (2005) Understanding GPS principles and applications, 2nd edn. Artech House Publishers, London, p 11
4. Recommendation ITU-R M.1583-1 (2002–2007) Interference calculations between non-geostationary mobile-satellite service or radionavigation-satellite service systems and radio astronomy telescope sites
5. Recommendation ITU-R RA.1631. Reference radio astronomy antenna pattern to be used for compatibility analyses between non-gso systems and radio astronomy service stations based on the epfd concept

Part II
Satellite Navigation Augmentation and Integrity Monitoring

Chapter 12
Modeling and Analysis for Availability of System-Level Integrity Monitoring of Satellite Navigation System

Zhuopeng Yang, Heng Zheng, Meijun Fan and Liucheng Chen

Abstract Bayesian Networks (BN) and Markov Chain are applied to modeling and analysis for availability of system-level integrity monitoring in this paper. Failure rate and repair rate are combined with Markov Chain for computing availability of equipments. According to availability of equipments and their relationships, availability and state availability of Monitoring Stations and Monitoring Areas are obtained. By outages analyzing and simulating of SISMA, the complicated relationship of Monitoring Stations and system-level integrity monitoring model are determined. By fault diagnosis and analysis, availability and weakness of system-level integrity monitoring are confirmed. Two new design schemes are proposed and compared for improving availability. Result show that this method can be used for availability analysis and fault diagnosis.

Keywords Integrity monitoring · Availability · Bayesian networks · Markov chain · Fault diagnosis

12.1 Introduction

Integrity is defined to be the trust which can be placed in the correctness of the information provided by signal-in-space [1]. Because position errors from satellites can only be eliminated by detecting and informing users early, integrity monitoring is a significant technology for normal operation of satellite navigation system.

Z. Yang (✉) · H. Zheng
China Academy of Aerospace Standardization and Product Assurance,
Beijing 100071, China
e-mail: yangzhuopeng88@163.Com

M. Fan · L. Chen
Beijing Global Information Center of Application and Exploitation,
Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_12,

Integrity monitoring has been widely applied in several years. The Galileo integrity concept is discussed in Oeiiler [2], the algorithm of SISMA (Signal in Space Monitoring Accuracy) is given in Blomenhofer and Zhao et al. [3, 4], the outage of GPS Monitoring Stations is analyzed in Stanton and Strother [5], and the distribution of Monitoring Stations based on coverage is simulated in document [6].

Monitoring Station, which has functions of data receiving, time matching, and integrity monitoring, is an important part of navigation system. By absorbing ideas of those documents, BN and Markov Chain are applied to modeling of system-level integrity monitoring based on Monitoring Stations in this paper, fault diagnosis is carried out, and two schemes are proposed for improving availability.

12.2 Modeling and Analysis

Integrity monitoring is influenced by equipments, availability and distribution of Monitoring Stations, and state availability of Monitoring Areas (constituted by several Monitoring Stations). There is a complicated network relationship among Monitoring Stations. BN [7] is applied in this paper to solve above relationship and multi-state. The concrete steps of modeling and analysis are as follows:

(a) Calculating the availability of equipments; (b) Calculating the availability of Monitoring Stations, and state-availability of Monitoring Areas; (c) Calculating the relationship of Monitoring Stations and Monitoring Areas, and model building; (d) Availability and sensitivity analysis and improved model.

The flow chart of model and analysis of system-level integrity monitoring is shown in Fig. 12.1.

12.3 Availability Model of Integrity Monitoring Based on BN

12.3.1 Availability model of Equipments, Monitoring Stations and Monitoring Areas

12.3.1.1 Availability Model of Equipments

Markov Chain is applied to solve normal-fault states and to calculate availability of equipments. Maintenance time of equipments are classified as short, middle and long time, corresponding μ_i (repair rate) and λ_i (fault rate), $i = 1, 2, 3$. States transition of equipment is shown in Fig. 12.2, where S_1 is normal state and S_0 is fault state.

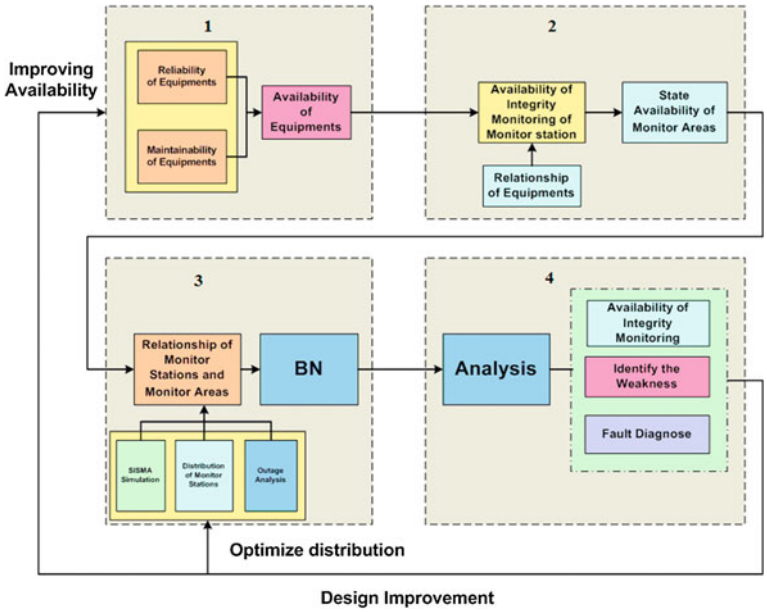


Fig. 12.1 The flow chart of model and analysis for availability of system-level integrity monitoring

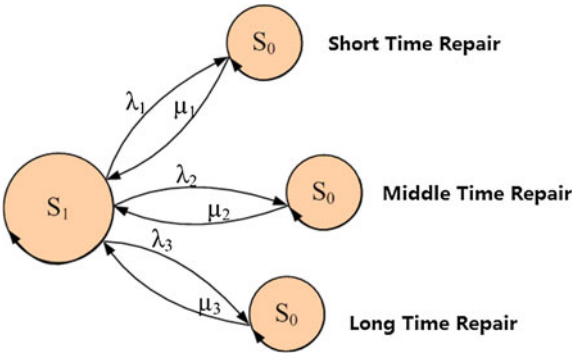


Fig. 12.2 Normal-fault states of an equipment by Markov chain

Fault rate and repair rate of an equipment are calculated as:

$$\begin{cases} \lambda = \lambda_1 + \lambda_2 + \lambda_3 \\ \mu = \dfrac{\lambda_1 + \lambda_2 + \lambda_3}{\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} + \frac{\lambda_3}{\mu_3}} \end{cases} \tag{12.1}$$

And the steady availability of an equipment is represented by:

$$A_e(\infty) = \frac{\mu}{\lambda + \mu} \tag{12.2}$$

12.3.1.2 Availability Model of Monitoring Stations

Monitoring Station is composed of many equipments (such as monitoring receivers, atomic clocks and data processing servers) and their relationships to perform its functions. Figure 12.3 shows the sketch map of Monitoring Station. The availability model of Monitoring Station by BN is illustrated in Fig. 12.4, and the relationship of monitoring receivers about 2/3-normal is shown in Fig. 12.4b.

12.3.1.3 State Availability Model of Monitoring Areas

26 Monitoring Stations are selected and named as M_i, i = 1,2,…,26. They are classified by transitional nodes as Northeast (Rne: $M_1 \sim M_5$), Southeast (Rse: $M_6 \sim M_{10}$), Middle Area (Rmp: $M_{11} \sim M_{15}$), Northwest (Rnw: $M_{16} \sim M_{20}$), Southwest (Rsw: $M_{21} \sim M_{26}$). Figure 12.5 shows the distribution of Monitoring Stations.

Assume that each Monitoring Area consists of m (m < 6) Monitoring Stations and n of m can be available, thus there is C_m^n combinations. Let available Monitoring Stations are $j_1, j_2, \ldots, j_n$, and $j_{n+1}, j_{n+2}, \ldots, j_m$ are unavailable. And $A_{j_i}, i = 1, 2, \ldots, m$ represent the availability of Monitor Station. Suppose that faults of Monitoring Stations are interrelated, state availability in k_{th} of Monitoring Area $P_{m,k}$ is therefore:

$$P_{m,k} = \left(\prod_{i=1}^{n} A_{j_i} \right) \cdot \left(\prod_{i=n+1}^{m} \left(1 - A_{j_i}\right) \right) \tag{12.3}$$

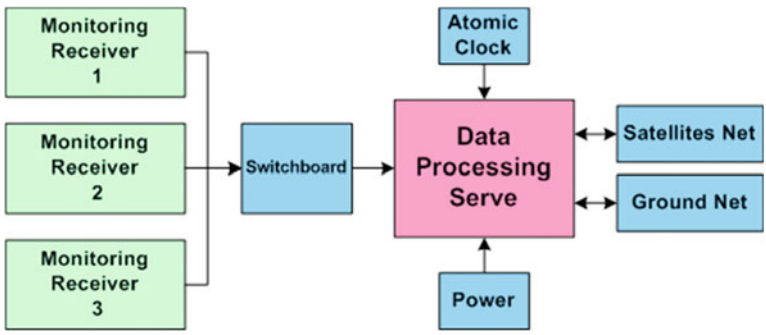
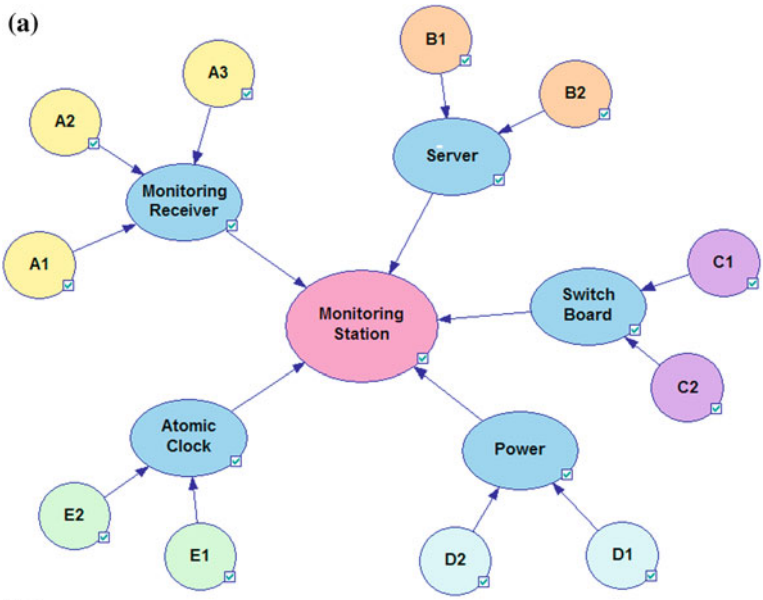


Fig. 12.3 Ketch map of monitoring station



A1	success				fault			
A2	success		fault		success		fault	
A3	success	fault	success	fault	success	fault	success	fault
success	1	1	1	0	1	0	0	0
fault	0	0	0	1	0	1	1	1

Fig. 12.4 Availability model of monitoring station based on BN. **a** BN of monitoring station. **b** CPT of monitoring receivers, *A* Monitoring receiver, *B* Data processing server, *C* Switch board, *D* Power, *E* Atomic clock

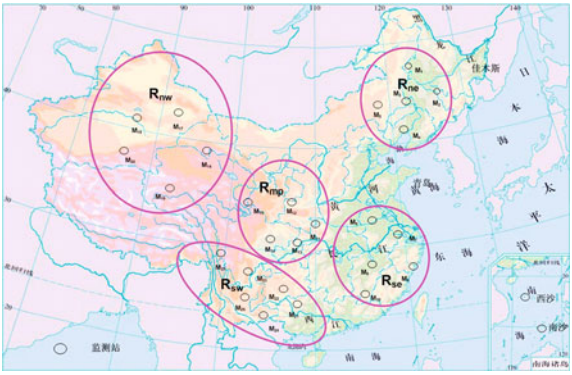


Fig. 12.5 The distribution of monitoring stations

12.3.2 Availability Model of System-Level Integrity Monitoring

12.3.2.1 Relationship of Monitoring Stations

Referring to Galileo Integrity concept, user integrity risk is related with DOP, SISA (Signal in Space Accuracy), and SISMA. Because SISA depends on accuracy of satellite orbit, SISMA is selected to be the relationship of Monitoring Stations.

Let x be the satellite true position and clock bias, at a given instant, each satellite j is seen by n Monitoring Stations, observation equations after being linearized from user can be expressed as:

$$\begin{aligned}\Delta\rho &= \begin{bmatrix}\Delta\rho_1^j\\ \Delta\rho_2^j\\ \dots\\ \Delta\rho_n^j\end{bmatrix} = \begin{bmatrix}e_{1x}^j & e_{1y}^j & e_{1z}^j & 1\\ e_{2x}^j & e_{2y}^j & e_{2z}^j & 1\\ \dots & \dots & \dots & \dots\\ e_{nx}^j & e_{ny}^j & e_{nz}^j & 1\end{bmatrix}\begin{bmatrix}\Delta x_s\\ \Delta y_s\\ \Delta z_s\\ \Delta t_s\end{bmatrix} + \begin{bmatrix}b_1^j\\ b_2^j\\ \dots\\ b_n^j\end{bmatrix} \\ &= Gx + b\end{aligned} \tag{12.4}$$

where $\Delta\rho$ is residual vector of the pseudorange observations from Monitoring Stations, G is observation matrix defining the direction cosine of sight vectors from user to the satellites, and b is observation noise.

Then the expressions of estimation and covariance of vector x, and error matrix of covariance σ_x^2 can be expressed as:

$$x = \left(G^T D_b^{-1} G\right)^{-1} G^T D_b^{-1} \Delta\rho \tag{12.5}$$

$$\sigma_x^2 = \left(G^T D_b^{-1} G\right)^{-1} \tag{12.6}$$

where D_b is pseudorange noise with a standard deviation in normal state.

So SISMA at user position can be computed as:

where $a_{uS} = (e_{uX}\ e_{uY}\ e_{uZ}\ 1)^T$ is the direction cosine vector from user to satellites.

$$\sigma_{SISEest}^2 = a_{uS}^T \sigma_x^2 a_{uS} \tag{12.7}$$

Thus, $SISMA = \max\left(\sigma_{SISEest}^2\right)$.

Simulating SISMA for 26 Monitoring Stations. Selecting a MEO which is on orbit inclination 55°, semi-major axis 27,906 km, and eccentricity ratio 0. Considering three states: all of Monitoring Stations are normal, any one of Monitoring Stations is operating outage, and any two of Monitoring Stations are operating outage simultaneously. Thus, three classes with 352 states are shown in Table 12.1. The result of SISMA simulation about the fist and the second class is represented in Fig. 12.6.

Table 12.1 Simulation condition of monitoring stations

No.	States	Times
1	All 26 monitoring stations are normal	1
2	Any one outages	26
3	Any two outage simultaneously	325

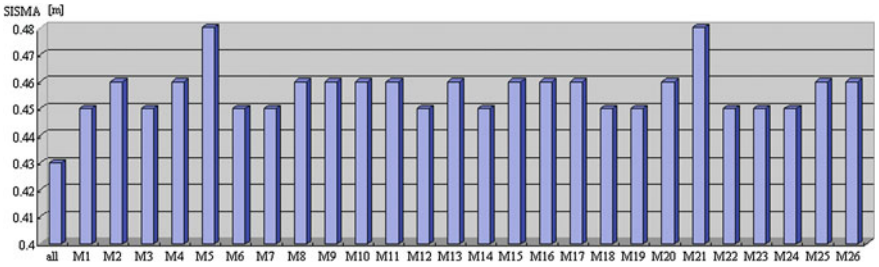


Fig. 12.6 SISMA simulation by monitoring station

Similarly, the third class can be simulated. In order to apply BN, $\sigma^2_{SISEest}$ should be transformed to $SISMA_{BN}$ by stress-strength method, and considered as the input CPT of BN. Selecting SISMA = 0.6 m as stress and $\sigma^2_{SISEest}$ as strength. Calculating $SISMA_{BN}$ by Monte-Carlo, selecting 70 % confidence and simulating 1,000 times:

$$SISMA_{BN} = P_r\left(E > \sum\right) = \int_0^\infty f_{\sum}(\sigma)d\sigma \int_0^\infty f_E(\varepsilon)d\varepsilon \tag{12.8}$$

where E is strength index, $\sum$ and is stress index. $SISMA_{BN}$ can be seen in the last column of Table 12.2.

Table 12.2 Conditional probability of monitoring stations

No.	State of Monitoring Stations																										SISMA	$SISMA_{BN}$
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26		
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.44	0.9999991
2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.48	0.99995599
3	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.45	0.99999732
4	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.45	0.99999732
5	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.45	0.99999732
6	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.46	0.99999351
7	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.48	0.99995599
8	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.45	0.99999732
9	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.45	0.99999732
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
28	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.49	0.99990919
29	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.51	0.99974118
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
325	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0.50	0.99973111

Note In the state of Monitoring Stations, *1* means normal, *0* means outage

12.3.2.2 Availability model of Integrity Monitoring

1. According to operational states of Monitoring Station, and data statistics during 365 days, fault rate, repair rate and steady availability are conformed. Suppose that reliability and maintainability firm exponential distribution. Data of equipments is shown in Table 12.3.
2. Combining with availability and relationship of equipments, steady availability of Monitoring Stations can be determined in Table 12.4.
3. Combing with classification of outages, state availability of Monitoring Areas can be determined in Table 12.5.
4. Determining the relationship of Monitoring Areas, transforming SISMA to $SISMA_{BN}$ in CPT, and modeling availability of integrity monitoring based on BN as Fig. 12.7.

In Fig. 12.7a, M1 ~ M26 are Monitoring Stations. Figure 12.7b is state availability of Monitoring Areas, and A0 ~ A3 represent the four states. Figure 12.7c is CPT of Monitoring Areas.

12.4 Analysis and Improvement of Availability by Redesign

Steady availability can be calculated by reasoning as Eq. (12.9).

$$\begin{aligned} &P\left(A, R_{ne}, R_{se}, R_{sw}, R_{mp}, R_{nw},_1, \ldots, M_{26}\right) \\ &= P\left(A \middle| R_{ne}, R_{se}, R_{sw}, R_{mp}, R_{nw}\right) \cdot P(R_{ne}, |M_1, \ldots, M_5) \cdot P(R_{se}, |M_6, \ldots, M_{10}) \\ &\cdot P(R_{sw}, |M_{11}, \ldots, M_{15}) \cdot P\left(R_{mp}, |M_{16}, \ldots, M_{20}\right) \cdot P(R_{nw}, |M_{21}, \ldots, M_{26}) \\ &\cdot P(M_1) \cdot P(M_2) \ldots \cdot P(M_{26}) \\ &= 0.9988 \end{aligned} \tag{12.9}$$

Table 12.3 Data of equipments

Equipments	Fault rate (10–6/h)	Repair rare (/h)	Number	Relationship[a]
Monitoring receiver	125	1	3	2/3
Data processing serve	38	0.5	2	1/2
Switch broad	11.5	1	2	1/2
Power	14	1.8	2	1/2
…	…	…	…	…

[a] n/m: at least n of m are normal then their combination are normal

Table 12.4 Steady availability of integrity monitoring of monitoring stations

No.	Availability	No.	Availability	No.	Availability	No.	Availability
M_1	0.9937	M_8	0.9969	M_{15}	0.9921	M_{22}	0.9983
M_2	0.9983	M_9	0.9948	M_{16}	0.9938	M_{23}	0.9923
M_3	0.9968	M_{10}	0.9948	M_{17}	0.9973	M_{24}	0.9902
M_4	0.9926	M_{11}	0.9965	M_{18}	0.9927	M_{25}	0.9995
M_5	0.9958	M_{12}	0.9967	M_{19}	0.9938	M_{26}	0.9998
M_6	0.9937	M_{13}	0.9982	M_{20}	0.9997		
M_7	0.9963	M_{14}	0.9918	M_{21}	0.9943		

Table 12.5 State availability of monitoring areas

State	State availability				
	Northeast	Southeast	Middle	Northwest	Southwest
All monitor station normal	0.97740	0.97672	0.97553	0.97749	0.97463
Any one outages	0.02241	0.02307	0.02425	0.02232	0.02514
Any two outage	0.00018	0.00020	0.00021	0.00019	0.00022
More than two outage	≈ 0	≈ 0	0.0001	≈ 0	≈ 0

Suppose that fault diagnosis is carried out under the fault state of integrity monitoring, then the weakness of Monitoring Areas and Stations is identified. Fault probability of Northeast is given by the following expression:

$$P(R_{ne} = 0|A = 0) = \frac{P(A = 0|R_{ne} = 0)P(R_{ne} = 0)}{P(A = 0)}. \tag{12.10}$$

Similarly, other Monitoring Areas can be identified in Fig. 12.8.

In Fig. 12.8, A1 means any one of Monitoring Stations outages, and so as A2, A3. It is known that Southwest area is the weakness of whole Monitoring Areas. Further analysis shows the number of Monitoring Stations in Southwest Area is more than other areas, and the availability of some Monitoring Stations is lower than other stations. All of those factors lead that Southwest area is weakness.

Furthermore, the weakness Monitoring Station ($M_{21} \sim M_{26}$) of Southwest area on fault state is analyzed, and is shown in Fig. 12.9.

Figure 12.9 shows that M_{24} is the weakness of Southwest area. The importance list of Southwest area is $M_{24} > M_{23} > M_{21} > M_{22} > M_{25} > M_{26}$. By analysis, it is known that the impaction of availability of Monitoring Stations is more important than their distribution when they are intensive in conforming weakness.

Two schemes for M_{24} are put forward and compared to improve the availability of integrity monitoring. The First scheme is improving availability of M_{24} by reducing fault rate by 5 % (such as Monitoring Receivers), and raising repair rate by 5 % (such as Data Processing Servers). Renew computing, the result shows the availability of M_{24} raising from 99.02 to 99.03 %, and system-level integrity monitoring raising from 99.88 to 99.89 %. The second scheme is adding another same level Monitoring Stations as M_{24}. Renew model, the result shows the

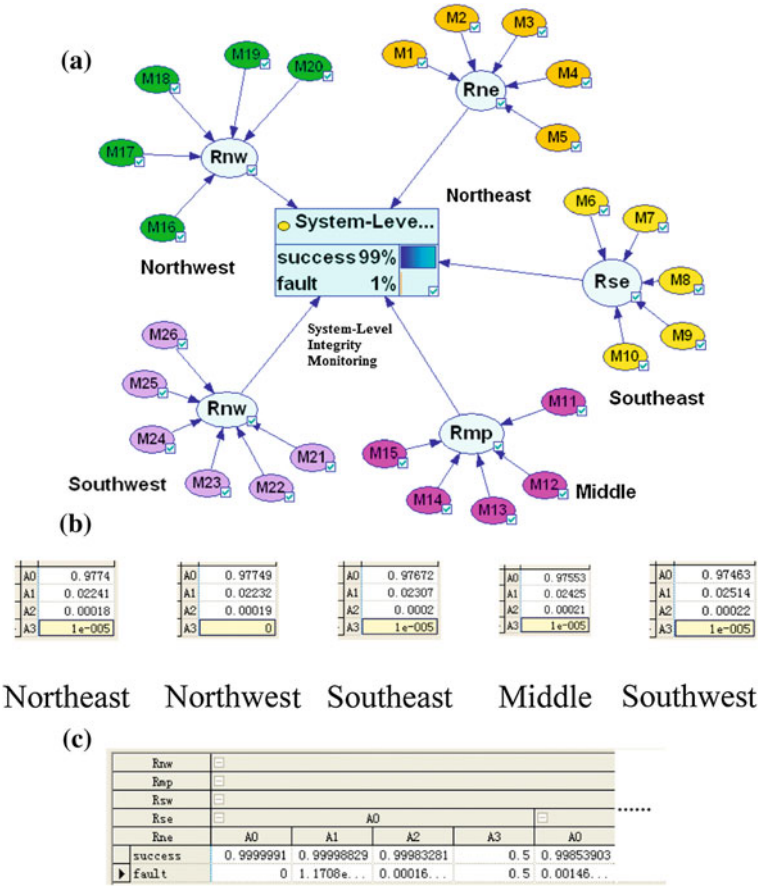
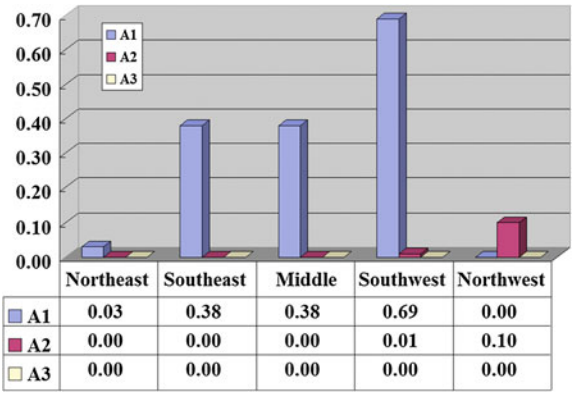


Fig. 12.7 Model of availability of system-level integrity monitoring. **a** BN of Integrity monitoring. **b** State availability of monitoring areas. **c** CPT of monitoring areas

Fig. 12.8 Fault probability of monitoring areas



	Northeast	Southeast	Middle	Southwest	Northwest
A1	0.03	0.38	0.38	0.69	0.00
A2	0.00	0.00	0.00	0.01	0.10
A3	0.00	0.00	0.00	0.00	0.00

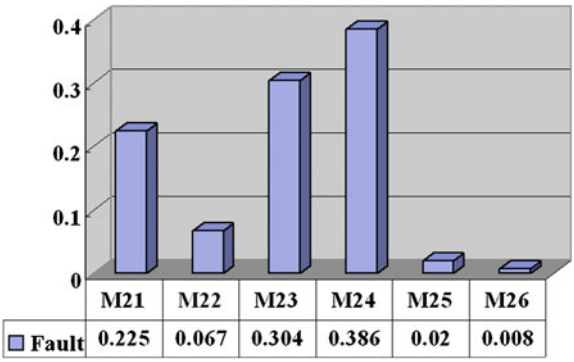


Fig. 12.9 Fault probability of Southwest areas

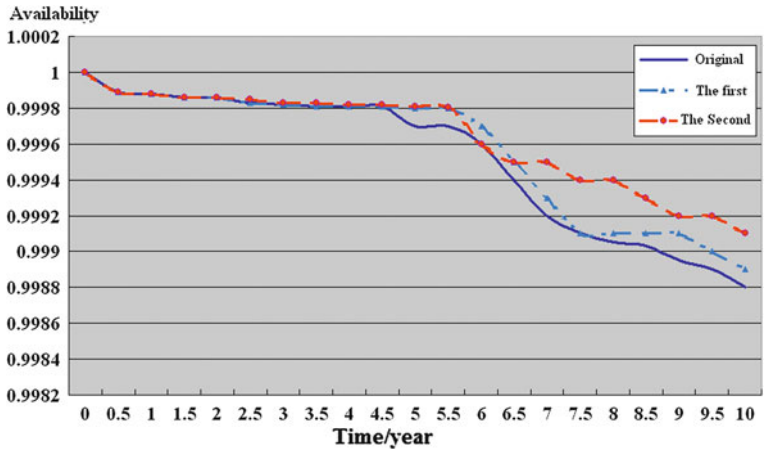


Fig. 12.10 Availability plots in different schemes

availability of system-level integrity monitoring raising from 99.88 to 99.91 %. Figure 12.10 shows availability plot of original, the first and the second scheme.

Compared to original scheme, it is shown that the first and the second scheme both are contributing to improve system-level integrity monitoring. Furthermore, due to an additional Monitoring Station in the second scheme, the relationship of Monitoring Stations (SISMA) is better than other schemes. The second scheme is better than the original and the first in fluctuating and value of availability, but needs more investments.

12.5 Conclusions

BN and Markov Chain have been applied together to modeling and analysis for integrity monitoring in this paper, and quantitative analysis and redesign is given.

(1) Quantitative analysis. According to repair time and fault rate, steady availability of equipments is determined by Markov Chain. The availability of Monitoring Stations is calculated based on BN. The complicated relationship of Monitoring Stations is obtained based on SISMA simulation, the model of system-level Integrity monitoring is constructed and the availability is 99.88 %.
(2) Confirming the weakness by fault diagnosis. The result shows, the importance list of Monitoring Areas is Southwest > Southeast = Middle > Northeast > Northwest, and importance list of Monitoring Stations in Southwest is $M_{24} > M_{23} > M_{21} > M_{22} > M_{25} > M_{26}$. Further analysis shows the impaction of availability of Monitoring Stations is more important than their distribution when they are intensive in conforming weakness.
(3) Redesigning and suggestion. Two schemes are approved for improving availability. The result shows: the availability of the first scheme is 99.89 %, and the second is 99.91 %. The second scheme is better than the original and the first in fluctuating and value of availability. Fault rate, repair rate, cost, distribution and strategy should be considered comprehensively on the scheme selection.

References

1. Department of Defense of United States of America (2008) Global positioning system standard positioning service performance standard. 4th Edition Sept 2008, 25–27
2. Oeiiler V et al. (2004) The Galileo integrity concept. In: Proceedings of ION GNSS 2004. Long Beach: ION, 604–615
3. Blomenhofer H (2005) Sensitivity analysis of the Galileo integrity performance dependent on the ground sensor station network. In: Proceeding of ION GPS/GNSS 18th international technical meeting of the satellite division, 1364–1365
4. Zhao C, Dang Y, Zhang H (2007) Calculation and analysis of integrity index SISMA of GALILEO system. ACTA GEODAETICA et CARTOGRAPHICA SINICA 36:2, 129–133
5. Stanton BJ, Strother R (2007) Analysis of GPS monitor station outages. In: Proceeding of ION GNSS 20th, ION, Texas, 176–183
6. Li Z, Hao J, Li J, Zhao Q, Zhang C (2010) Primary study on coverage and distributing methods of satellite navigation system monitoring stations. China satellite navigation conference 2010. Beijing: Organizing committee of China satellite navigation conference
7. Zheng H, Zhou H (2011) Probability risk assessment. National Defense Industry Press, Beijing, 8

Chapter 13
Minimum of GDOP of Satellite Navigation and Its Applications in ISL Establishment of Walker-δ Constellation

Songhui Han, Po Chen, Qingming Gui, Jianwen Li and Meng Wei

Abstract GDOP is one of the important standards which measure the accuracy of a navigation positioning system. Firstly, this article discusses the shortcomings in the domestic and foreign literatures of deriving the minimum of GDOP. The existing theories of minimum of GDOP are all based on the observations of ground stations, and the deduced GDOP minimum value cannot be reached because the observable range of the receiver do not be considered. Then, the minimum GDOP is derived under the condition which the cutoff angle of satellite is greater than zero, and a method to get the minimum GDOP is proposed and the corresponding geometry between satellites is given. This new theory of minimum of GDOP is also applied to the criteria for inter-satellite link establishment in Walker-δ constellation which is proposed by the authors. The optimal structure diagram of inter-satellite link is given in terms of GDOP. Finally, computations indicate the effective and useful of this new theory of the minimum of GDOP.

Keywords Satellite navigation · Minimum of GDOP · Matrix analysis · ISL

S. Han (✉)
Institute of Geospatial Information, Information Engineering University, Zhengzhou 450052, China
e-mail: hansonghui@126.com

P. Chen · J. Li
Institute of Navigation and Aerospace Engineering, Information Engineering University, Zhengzhou 450052, China

S. Han · Q. Gui · M. Wei
Institute of Science, Information Engineering University, Zhengzhou 450001, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_13,

13.1 Introduction

The accuracy of navigation positioning of GNSS are determined not only by the accuracy of observations, but also by the geometry of receivers and satellites. GDOP (Geometric Dilution of Precision) describes this geometry and the geometrical effect on the proportion of measurement error of pseudorange to position error of target, and it is an important index to measure the positioning accuracy [1]. The better the geometry, the smaller the value of GDOP, and the higher the positioning accuracy can be got. So far, most of the related researches of GDOP are based on the satellite-ground observations. Jwo and Lai [2] present a selection method of navigation satellite subset to evaluate all subsets of satellites, and this approach is based on approximation or classification of the satellite GDOP factors by means of the neural network approach. In order to eliminate the influence of obstructions on GPS signals and forecast the positioning accuracy of GPS, Lohani and Kumar [3] attempt to develop a model which uses LiDAR data to predict GDOP at a point in space and time, and this model can estimate GPS satellite visibility and its role in GPS positional accuracy by using GDOP. In all the above researches, the minimum of GDOP should be obtained so as to judge whether the approximate or predicted GDOP is correct. In addition, in order to improve the positioning accuracy of GNSS, the minimum of GDOP should also be discussed [3, 4]. Cong et al. [5], Zhang et al. [6] and Cong and Tan [7] introduce the relation between GDOP and volume of polyhedron composed of receiver and satellites, but they do not give the minimum of GDOP. Yang and Miao [8] analyze the GDOP of positioning of GPS and deduce the result that the minimum of GDOP is $\sqrt{2}$ when 4 satellites are observed. However, Sairo et al. [9] point out that the minimum of GDOP cannot reach $\sqrt{2}$, and they use matrix eigenvalue to give a new minimum of GDOP, which is $\sqrt{10/n}$. This new minimum is also obtained at a different angle by Sheng et al. [1] and Li et al. [10]. All the above results are useful for the discussion of the minimum of GDOP. But an analysis shows that $\sqrt{10/n}$ is obtained in terms of pure mathematics, neglecting the fact that receiver can not receive the signals when satellites are under the horizon. The minimum of GDOP cannot reach $\sqrt{10/n}$ in satellite positioning, for receiver can only receive the signals when the elevation mask angle of satellite is greater than zero [11].

Since Ananda et al. [12] proposed the idea of autonomous orbit determination, it has made tremendous development. But the Earth Orientation Parameters (EOP) as one of the error sources cannot be determined when only inter-satellite observations are used. In fact, the observations that connect the satellite orbits to the geocentric inertial coordinate system must be provided. The commonly used methods of providing these observations include: fitting orientation parameters through experience or the application of prior information [13], ground station observations [14, 15], and star or pulsar observations [4]. The mode based on inter-satellite measurements combined with these observations is a feasible scheme for autonomous orbit determination, and reasonable inter-satellite links (ISLs) should

be designed for this mode. In addition, one can assume that a future satellite navigation system does not only provide a navigation service, but also a communication payload enabling global communication by using ISLs [16].

The paper is organized as follows: Firstly, the shortcomings of the methods of deriving the minimum of GDOP in domestic and foreign literatures are analyzed in Sect. 13.2. The new minimum GDOP is derived under the condition that the elevation mask angle of satellite is greater than zero and the corresponding satellites geometry is given in Sect. 13.3. In Sect. 13.4, we use the new theory of minimum of GDOP to verify the existing criteria for ISL establishment in Walker-δ constellation. In order to demonstrate the feasibility and usefulness of the new theory, two examples for the GDOP of satellite-ground observations and inter-satellite observations are given in Sect. 13.5. Finally, some concluding remarks are given at the end of the paper.

13.2 The Existing Theories of Minimum of GDOP

If the receiver continuously observes the same $n(n \geq 4)$ satellites, through linearization of observation equations around a priori receiver position $\vec{r}_{uo} = (x_0\ y_0\ z_0)^T$, in matrix form, observations equation can be written as:

$$HX = L + \varepsilon \tag{13.1}$$

where: $H = \begin{bmatrix} e_{x1} & e_{y1} & e_{z1} & 1 \\ e_{x2} & e_{y2} & e_{z2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ e_{xn} & e_{yn} & e_{zn} & 1 \end{bmatrix}$ is the $n \times 4$ design matrix, $e_i = (e_{xi}, e_{yi}, e_{zi})$ is the direction cosine of receiver pointing to the ith satellite, and $|e_i| = 1$, that is to say, $e_i (i = 1, 2, \ldots, n)$ are all on unit sphere; the n columns of H are linearly independent since they are signals received from individual satellites independently; $X = (\delta x\ \delta y\ \delta z\ b)^T$ is the unknown parameters vector, $(\delta x\ \delta y\ \delta z)^T$ is the vector of corrections and b is the equivalent distance parameter of the clock error of receiver; L is an $n \times 1$ vector of observations and the weight matrix is unit weight; ε is an $n \times 1$ vector of observation noises.

The least squares estimator of unknown parameters vector X in model (13.1) can be expressed as

$$\hat{X} = (H^T H)^{-1} H^T L \tag{13.2}$$

The covariance matrix of least squares estimator is $Q = \sigma_0^2 (H^T H)^{-1}$, where σ_0^2 is an unknown variance component of unit weight. Thus, the GDOP factor is defined as $GDOP = \sqrt{\mathrm{tr}(H^T H)^{-1}}$, where $\mathrm{tr}(H^T H)^{-1}$ is the trace of matrix $(H^T H)^{-1}$.

The space distribution feature of receiver and visual satellites has been described by GDOP. The best space geometry is obtained when GDOP reaches the minimum value, and the regional navigation performance is excellent. Otherwise, the regional navigation performance will be bad or regional navigation can not be used. In addition, the minimum of GDOP decreases gradually with the increase of the number of visual satellites. The greater the number of visual satellites, the smaller the value of GDOP is [10, 11].

Assume $\lambda_i (i = 1, 2, 3, 4)$ are the eigenvalues of the matrix $H^T H$, then

$$GDOP = \sqrt{\sum_{i=1}^{4} \frac{1}{\lambda_i}} \tag{13.3}$$

In order to get the minimum of GDOP, the following inequality has been proved; see e.g., Sheng et al. [1] or Li et al. [10]

$$GDOP = \sqrt{\sum_{i=1}^{4} \frac{1}{\lambda_i}} = \sqrt{\sum_{i=1}^{3} \frac{1}{\lambda_i} + \frac{1}{\lambda_4}} \geq \sqrt{3\left[\prod_{i=1}^{3} \frac{1}{\lambda_i}\right]^{\frac{1}{3}} + \frac{1}{\lambda_4}} \tag{13.4}$$

Then, using the equality $|e_i| = 1, i = 1, 2, \ldots n$, the trace of matrix $H^T H$ becomes

$$\mathrm{tr}\left(H^T H\right) = 2n \tag{13.5}$$

that is

$$\sum_{i=1}^{4} \lambda_i = 2n \tag{13.6}$$

The equality sign of (13.4) holds only if $\lambda_1 = \lambda_2 = \lambda_3$, and the result $GDOP_{\min} = \sqrt{\frac{10}{n}}$ which is the minimum of GDOP can be got when $\lambda_4 = n$. The same minimum of GDOP can also be got when $\sum_{i=1}^{n} e_{xi}^2 = \sum_{i=1}^{n} e_{yi}^2 = \sum_{i=1}^{n} e_{zi}^2 = \frac{n}{3}$ [11]. This minimum of GDOP was made based on the premise that the matrix $H^T H$ is a diagonal matrix, that is, the elements of matrix H must meet the following conditions:

$$\begin{cases} \sum_{i=1}^{n} e_{xi} e_{yi} = \sum_{i=1}^{n} e_{xi} e_{zi} = \sum_{i=1}^{n} e_{yi} e_{zi} = 0 \\ \sum_{i=1}^{n} e_{xi} = \sum_{i=1}^{n} e_{yi} = \sum_{i=1}^{n} e_{zi} = 0 \end{cases} \tag{13.7}$$

Because the satellites involved in positioning calculation must be all visual satellites over the ground plane and the elevation mask angle of satellite is $(0, \frac{\pi}{2})$, $e_{zi} > 0 \ (i = 1, 2, \ldots, n)$ should be satisfied. This is in contradiction with the

condition $\sum_{i=1}^{n} e_{zi} = 0$. Thus, $H^T H$ can not be a diagonal matrix, and $GDOP_{\min} = \sqrt{\frac{10}{n}}$ cannot be reached when receiver is on the earth's surface.

13.3 Analysis of Minimum of GDOP

When receiver is on the earth's surface, the derivation of minimum of GDOP must consider the condition $e_{zi} > 0$ $(i = 1, 2, \ldots, n)$. After analysis, the following conclusion is obtained:

Theorem *Assume* $H^T H = \begin{bmatrix} R & U \\ U^T & T \end{bmatrix}$, *where*

$$R = \begin{bmatrix} \sum_{i=1}^{n} e_{xi}^2 & \sum_{i=1}^{n} e_{xi} e_{yi} \\ \sum_{i=1}^{n} e_{yi} e_{xi} & \sum_{i=1}^{n} e_{yi}^2 \end{bmatrix}, T = \begin{bmatrix} \sum_{i=1}^{n} e_{zi}^2 & \sum_{i=1}^{n} e_{zi} \\ \sum_{i=1}^{n} e_{zi} & n \end{bmatrix}, U = \begin{bmatrix} \sum_{i=1}^{n} e_{xi} e_{zi} & \sum_{i=1}^{n} e_{xi} \\ \sum_{i=1}^{n} e_{yi} e_{zi} & \sum_{i=1}^{n} e_{yi} \end{bmatrix}$$

then $GDOP \geq \sqrt{\mathrm{tr}(R^{-1}) + \mathrm{tr}(T^{-1})}$, *and the equality sign is true if and only if* $U = 0$, *that is* $\mathrm{GDOP}_{\min} = \sqrt{\mathrm{tr}(R^{-1}) + \mathrm{tr}(T^{-1})}$.

$U = 0$ requires that the following conditions hold:

$$\sum_{i=1}^{n} e_{xi} e_{zi} = \sum_{i=1}^{n} e_{xi} = \sum_{i=1}^{n} e_{yi} e_{zi} = \sum_{i=1}^{n} e_{yi} = 0 \tag{13.8}$$

This transforms the minimum of GDOP problem into an equation set problem which is to find the solution of Eq. (13.8). However, it is difficult to get the analytic solution of Eq. (13.8). It is well known that the more dispersed the spatial distribution of satellites, the smaller the value of GDOP is. According to this fact, a constructive solution of Eq. (13.8) is given when $n = 4$, that is,

$$(e_x, e_y, e_z) = \begin{bmatrix} \sin\alpha & 0 & \cos\alpha \\ 0 & \sin\beta & \cos\beta \\ -\sin\alpha & 0 & \cos\alpha \\ 0 & -\sin\beta & \cos\beta \end{bmatrix}, \quad 0 < \alpha, \beta < \frac{\pi}{2}, \quad \alpha \neq \beta \tag{13.9}$$

After analysis, it can be found that this solution does not only satisfy Eq. (13.8), but also meets the condition $\sum_{i=1}^{n} e_{xi} e_{yi} = 0$. Thus, matrix R becomes a diagonal matrix. The clear and definite geometric meaning of this solution is as follows: (1) the positions of the receiver and 4 satellites compose a symmetrical hexahedron; (2) the central axis of this hexahedron coincides with the Z axis which points to center of the earth, and the position of receiver is in the Z axis; (3) the 4 satellites

are not on a plane which parallels the plane XOY, and the projection of this hexahedron onto the plane XOY is a rhombus. If $\alpha = \beta$ is satisfied in Eq. (13.9), then 4 satellites will be all on a plane which parallels the plane XOY, and the receiver and 4 satellites compose a pentahedron; but the coefficient matrix H is singular, that is to say, the solution obtained when $\alpha = \beta$ in Eq. (13.9) cannot meet the requirement of positioning.

The expression of the minimum of GDOP when $n = 4$ can be got as follows by putting the solution (13.9) into matrix $H^T H$

$$GDOP_{\min} = \sqrt{\frac{\cos^2\alpha - \cos^2\beta + 2}{2(\cos\alpha - \cos\beta)^2} + \frac{1}{2\sin^2\alpha} + \frac{1}{2\sin^2\beta}} \tag{13.10}$$

With the numerical method, the minimum of GDOP can be readily obtained from (13.10).

Similar to the above, a constructive solution of Eq. (13.8) is given when $n = 5$

$$(e_x, e_y, e_z) = \begin{bmatrix} \sin\alpha & 0 & \cos\alpha \\ 0 & \sin\beta & \cos\beta \\ -\sin\alpha & 0 & \cos\alpha \\ 0 & -\sin\beta & \cos\beta \\ 0 & 0 & 1 \end{bmatrix}, \quad 0 < \alpha, \beta < \frac{\pi}{2}, \quad \alpha \neq \beta \tag{13.11}$$

Then, the corresponding expression of minimum of GDOP is as follows

$$GDOP_{\min} = \sqrt{\frac{\cos^2\alpha + \cos^2\beta + 3}{3\cos^2\beta + 3\cos^2\alpha - 4\cos\alpha\cos\beta - 2\cos\alpha - 2\cos\beta + 2} + \frac{1}{2\sin^2\alpha} + \frac{1}{2\sin^2\beta}} \tag{13.12}$$

Based on the geometry of $n = 4$, the position of the fifth satellite is also in the Z axis, and the minimum of GDOP can be got from (13.12).

In the same way, the solutions which satisfy $U = 0$ and $\sum_{i=1}^{n} e_{xi} e_{yi} = 0$ can be constructed when $n = 2k$ or $n = 2k + 1$, and the minimum of GDOP can be obtained by using the numerical method.

13.4 The Establishment of ISL in Walker-δ Constellation

Each MEO satellite has the same status in the Walker-δ constellation, which belongs in distributed topology and is symmetric. GPS, GLONASS and Galileo satellite navigation systems all use Walker-δ constellation with different numbers of orbital planes. GPS has 6 orbital planes, but it is a custom Walker 24/6/1 constellation configuration rather than a strictly standard Walker-δ one. All satellites in Walker-δ constellation have the same orbit altitude, and thus all satellites

move on the surface of an approximate sphere. If we construct a plane which is perpendicular to the line connecting the centre of the earth with one satellite, the satellite being on this plane, then all other satellites are located on the same side of this plane and the ISLs between this satellite and all other satellites are also on this side. That is, one satellite can only establish ISLs with satellites on one side of this plane, just as receivers can only see satellites over the ground plane. Thus, the theory of minimum of GDOP can be used to discuss the establishment of ISL in Walker-δ constellation, and ISLs which meet the requirements of autonomous orbit determination can be established by using the conclusion of this paper.

The establishment of ISL is mainly affected by the azimuth angle, elevation angle and transmission distance between satellites. The smaller the ranges of azimuth angle, elevation angle and transmission distance of ISL, the better the performance of satellite transmitting information by using ISL will be. Thus, under the requirements of short average transmission distance and small ranges of the azimuth angle, elevation angle and transmission distance of ISL, the criteria for ISL establishment in Walker-δ sub-constellation are given as follows [17, 18]:

1. In the same orbital plane, establish ISLs between adjacent satellites.
2. Among the adjacent orbital planes, establish ISLs between satellites with the shortest average transmission distance and the smallest ranges of the azimuth angle, elevation angle and transmission distance of ISL.
3. The azimuth angle and elevation angle of the ISLs in the same orbital plane should be as much as possible contained in the ranges of azimuth angle and elevation angle of the ISLs among the adjacent orbital planes.

The third criterion is proposed considering the visual range of single beam antennas. According to the above criteria, each MEO satellite in Walker-δ constellation can establish four ISLs with other MEO satellites, of which two ISLs are built between the two adjacent MEO satellites in the same orbital plane respectively and the other two could be established between MEO satellites in the adjacent sides of this orbital plane respectively. These ISLs are all permanent, and the entire constellation is 4-connected network ("4-connected" is a term of graph theory); for more details, see Han et al. [17] and [18].

In order to improve the geometry structure of inter-satellite observations and to facilitate autonomous orbit determination, the fourth criterion is proposed as follows according to the theory of minimum of GDOP:

4. The corresponding GDOP value of inter-satellite observations should be close to the minimum when each satellite establishes ISL with other satellites.

If satellite antenna belongs in the single beam antennas, in order to save the limited on-board resources, the third criterion is more important compared with the fourth criterion. But, the third criterion can be neglected when omni directional antenna has been adopted.

13.5 Simulation and Analysis

In this section, we will use two synthetic examples to demonstrate the theory of minimum of GDOP proposed in this paper.

Example 1 In order to verify the correctness of the minimum of GDOP deduced in this paper, simulation and analysis is given by taking 4 visible satellites and 5 visible satellites for example. The following 3 schemes are carried out to compute the corresponding minimum of GDOP. The calculation results are listed in Table 13.1.

1. The theory of minimum of GDOP proposed in this paper, denoted as $\mathrm{GDOP}_{\mathrm{min1}}$.
2. The minimum of GDOP computed by $GDOP_{\min} = \sqrt{\frac{10}{n}}$, denoted as $\mathrm{GDOP}_{\mathrm{min2}}$.
3. According to the orbital elements of satellite provided by the website of Jet Propulsion Laboratory, without considering the influence of various errors, the minimum of GDOP is computed by extrapolation method. This minimum is relatively accurate, denoted as $\mathrm{GDOP}_{\mathrm{true}}$.

From the results in Table 13.1, it can be found that the minimum of GDOP proposed in this paper is basically identical with the $\mathrm{GDOP}_{\mathrm{true}}$, and the minimum of GDOP cannot reach $\sqrt{\frac{10}{n}}$ because the restriction of the elevation mask angle of satellite is not considered.

Example 2 STK (Satellite Tool Kit) is used to simulate the constellation of Walker 24/3/1, and GPS SPS (2001) standards are used to determine the satellite altitude and orbital inclination parameters, which are respectively 26,559.8 km and 55°. Taking satellite A1 as example, GDOP value is discussed under the mode of permanent ISL. A new coordinate is established by taking the position of the satellite A1 as the origin: Z axis points to center of the earth, the X-axis points to the tangential direction of the direction of the satellite movement, and the Y-axis is perpendicular to the XZ plane. Considering the symmetry of Walker-δ constellation, the phase angles only between satellite A1 and satellites B1-B8 of the second orbital plane are considered. Because of the shade of the earth, satellite A1 can only be consistently visible to 6 satellites of B1 and B4-B8. The ISLs between satellite A1 and other 6 satellites of B1 and B4-B8 have been established, and the corresponding differences of phase angle are 15°, 150°, 195°, 240°, 285° and 330°. In one running period, the inter-satellite measurements are obtained by taking the

Table 13.1 Min GDOP value of three schemes

Number of visible satellites	$\mathrm{GDOP}_{\mathrm{min1}}$	$\mathrm{GDOP}_{\mathrm{min2}}$	$\mathrm{GDOP}_{\mathrm{true}}$
4	1.979603	1.5811	1.979799
5	1.581895	1.4142	1.581987

Table 13.2 Min GDOP value when A1 is the observation point

Differences of phase angle	15°	150°	195°	240°	285°	330°
$GDOP_{min}$	3.6594	3.7201	5.4991	3.7007	5.6280	3.3189

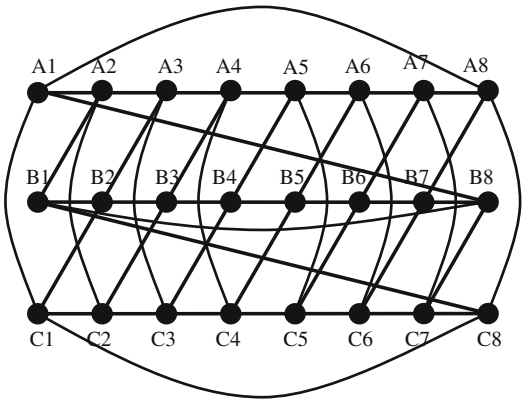


Fig. 13.1 ISLs based on the ISL establishment criteria of 1, 2 and 3

position of satellite A1 as the observational point, and then the minimum of GDOP are given in Table 13.2.

Table 13.2 shows that 3.3189 is the minimum of GDOP when ISL is established between satellites A1 and B8, that is to say, the difference of phase angle is 330°. According to the above criteria 1-3 of ISL establishment, ISLs are built between adjacent satellites in the same orbital plane, and between satellites in the adjacent orbital planes with the mean anomaly difference of 30°. There are 48 permanent ISLs in Walker 24/3/1 constellation, and all the permanent ISLs can be observed in Fig. 13.1. After analysis, the structure of ISLs which is based on establishment criteria 1, 2 and 4 are the same as in Fig. 13.1. Thus, the structure of ISLs according to criteria 1-3 is the one with the optimal GDOP value.

13.6 Conclusions

In this paper, the method of minimum of GDOP in GNSS positioning is analyzed, and some conclusions are made as follows:

1. In navigation and positioning, receiver can only receive signals of satellites over the ground plane, and some result of minimum of GDOP is unreasonable and the minimum cannot be reached.
2. The new minimum GDOP is derived, and the minimum of GDOP is given when the space geometry of satellites fulfils some conditions. In addition, the specific minimum of GDOP is obtained by using numerical method when the numbers of visual satellites are 4 and 5.

3. With the development of GNSS, the number of visual satellites will increase and the geometry of satellites will improve, and thus the value of GDOP will further decrease.
4. The new theory of minimum of GDOP is also applicable to the autonomous orbit determination. The optimal structure of ISLs can be established by using minimum of GDOP, and it helps to determine orbit of satellite.
5. The conclusions of this paper can be used as a reference for the design of new constellation configurations, accuracy analysis of navigation and positioning, establishment of ISL, and autonomous orbit determination.

Acknowledgments This research was supported jointly by National Science Foundation of China (No. 40974009, No.41174005), Planned Research Project of Technology of Zhengzhou City, and Funded Project with youth of Annual Meeting of China's satellite navigation, PhD Dissertation Innovation Fund of Information Engineering University.

References

1. Sheng Z, Yang JS, Zeng FL (2009) The minimum value of GDOP in pseudo-range positioning. Fire Control Command Control 34(5):22–24
2. Jwo DH, Lai CC (2007) Neural network-based GPS GDOP approximation and classification. GPS Solution 11(1):51–60
3. Lohani G, Kumar R (2007) A model for GPS-GDOP prediction in urban environment using LiDAR data. In: Proceedings of International conference on optical 3D measurement techniques, pp 339–346
4. Chester TJ, Butman SA (1981) Navigation using X-Ray pulsars.NASA TR N81-27129: 22–25
5. Cong L, Abidat AI, Tan ZZ (2006) Analysis and simulation of the GDOP of satellite navigation. Acta Electronica Sinica 34(12):2204–2208
6. Zhang YL, Fan L, Zhang Y et al (2008) Theory and design of satellite constellations. Science Press, Beijing
7. Cong L, Tan ZZ (2008) Satellite selection algorithm to improve precision and real-time performance of GPS positioning. Syst Eng Electron 30(10):1914–1917
8. Yang Y, Miao LJ (2004) GDOP results in all-in-view positioning and in four optimum satellites positioning with GPS PRN codes ranging. IEEE position location and navigation symposium, pp 723–727
9. Sairo H, Akopian D, Takala J (2003) Weighted dilution of precision as quality measure in satellite positioning. IEEE Radar, Sonar Navig 150(6):430–436
10. Li JW, Li ZH, Zhou W et al (2011) Study on the minimum of GDOP in satellite navigation and its applications. Acta Geodsetica at Cartographica Sinica 40:85–89
11. He XF (2009) Algorithms for BD/MIMU integrated navigation systems. National University of Defense Technology, Changsha
12. Ananda MP, Berstein H, Bruce RW et al (1984) Autonomous navigation of the global positioning system satellite. AIAA guidance and control conference, Seattle, WA, August, pp 20–22
13. Hou F, Liu WK, Fang RX (2009) Analysis and control method for constellation rotationin navigation satellite autonomous orbit determination. Geomatics Inf Sci Wuhan Univ 34(11):1280–1284

14. Rajan JA, Brodie P, Rawicz H et al (2003) Modernizing GPS autonomous navigation with anchor capability. GNSS 2003, September, Portland, pp 1534–1542
15. Zhu J, Wen YL, Chen ZG, Liao Y, Pan WH (2008) Research on modeling and simulation of semi-autonomous orbit determination for satellite navigation constellation. In: Proceedings of system simulation and scientific computing, Oct 2008, pp 252–257
16. Hermenier R, Kissling C, Donner A (2009) A delay model for satellite constellation networks with inter-satellite links. In: Proceedings of the international workshop on satellite and space communications, Siena-Tuscany, Italy, pp 3–7
17. Han SH, Gui QM, Li JW, Du YL (2011) Analysis of the connectivity and robustness of inter-satellite links in a constellation. Sci China Phys, Mech Astron 54(6):1–5
18. Han SH, Gui QM, Li JW, Chen Po, Du YL (2012) Analysis of the routing algorithm and using probability of inter-satellite link in Walker-δ constellation. China satellite navigation conference (CSNC) 2012 proceedings, lecture notes in electrical engineering vol 161, pp 333–341

Chapter 14
Ionospheric Spatial Correlation Analysis for China Area

Dun Liu, Liang Chen and Weimin Zhen

Abstract Ionospheric correlation models are needed for ionospheric grid-based augmentation system development and its rigorous integrity performance requirement. 0th and 1st order ionospheric correlation function was established for China area, and analyzed for various ionospheric conditions. It is shown that the ionospheric delay difference does not vary linearly with the distance, and the residual ionospheric error after the planar fitting of grid model is not a constant as well. These models are different from that of WAAS or EGNOS, which covers the middle latitude area. These China area ionosphere specific characteristics should be taken into account for the regional augmentation system development.

Keywords Integrity · Ionosphere · Anomaly · Storm · Spatial correlation model

14.1 Overview

Satellite Regional augmentation system generally employs ionospheric grid model for correction of ionospheric errors and relevant integrity realization in the system. In ionospheric grid model, a planar fitting is made with the measured ionospheric delay to acquire ionospheric delay estimation and relevant residual threshold estimation at the grid points, as well as users. Ionospheric delay is a spatial variable, therefore, characteristics of ionospheric delay spatial variation needs to be described in the ionospheric grid model.

The significance of ionospheric spatial correlation structure analysis lies in establishing a suitable ionospheric grid model and enhancing its accuracy. Regarding ionospheric correlation structure gives rise to changes of ionospheric

D. Liu (✉) · L. Chen · W. Zhen
China Research Institute of Radiowave Propagation, Qingdao 266107, China
e-mail: dun.l@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_14,

errors, a rational threshold shall be set for integrity of the system. Consequently, ionospheric spatial correlation structure analysis would give impact on the accuracy and integrity of the system.

Three types of Ionospheric spatial correlation model are needed, and they are:

1. 0th order spatial correlation model [1, 2]

0th order spatial correlation model of ionospheric delay describes correlation structure of ionospheric delay in the spatial distribution, which is mainly used in construction of regional ionospheric delay distribution.

2. 1st order spatial correlation model [2, 3]

1st order spatial correlation model is mainly used to describe, after planar fitting for ionospheric delay, changes of ionospheric residuals with the space. It is primarily used to set up a rational GIVE threshold and to realize ionospheric irregularities detection as well.

3. ionospheric spatial threat model [4, 5]

The description of effect for ionospheric residual threshold, due to "undersample" of ionospheric observation, will be explained in another paper.

In GPS WAAS system, perfect models said above have been established [1–5]. As WAAS primarily covers the middle latitude area of North America, while augmentation system in China principally covers middle and low latitude of Asia–Pacific area, whose characteristics of ionosphere changes are obviously different from that in middle latitude area of North America [6]. Hence, with a view to environment characteristics of ionosphere in China area, ionospheric spatial correlation models studies shall be carried out.

14.2 Data for Ionospheric Spatial Correlation Structure Analysis in China Area

14.2.1 GNSS Data

Data used in studies on ionospheric spatial correlation characteristics for China area mainly include:

1. GNSS dual-frequency observation data: the data from GPS observation stations (see Fig. 14.1) in China and the area around are used to construct regional ionospheric delay distribution.
2. Scintillation data: local disturbance aroused by scintillation is hard to be described by ionospheric spatial correlation model. Local ionospheric disturbance resulting from scintillation could be recognized and relevant data would be removed during the scintillation.

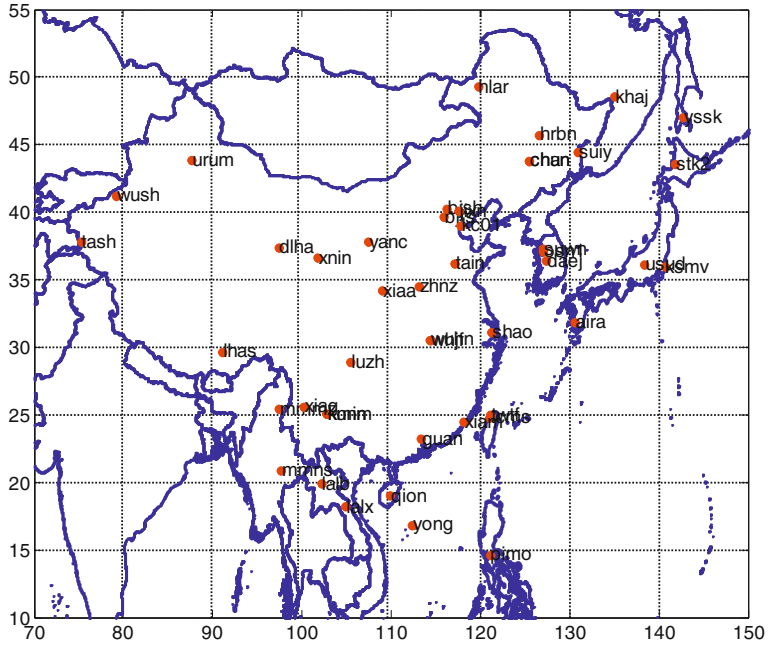


Fig. 14.1 Data analysis from GPS reference stations distribution in the low latitude area of China

3. Solar-Earth space environment data: ionospheric disturbance such as storm could be judged by making use of solar-earth space environment data issued by the international.

Time of data said above primarily covers different periods of solar activities and different ionospheric conditions (quiet condition, disturbance).

14.2.2 Pretreatment of GNSS Data

Ionospheric delay data, retrieved from GPS dual-frequency observation data, will adopt the following methods to make sure accuracy of ionospheric delay extracted.

1. Ionospheric delay is extracted by making use of carrier smoothed pseudo-range.
2. In the estimation of GPS hardware bias, DCB value of GPS satellite issued by CODE is introduced to enhance accuracy of receiver hardware bias estimation.
3. The receiver hardware estimation is averaged for several days' data to improve stability of hardware bias estimation, particularly for the period of ionospheric disturbance.

14.3 Ionospheric 0th Order Spatial Correlation Model

14.3.1 Method of Establishing Ionospheric 0th Order Spatial Correlation Model

As for implementation of ionospheric grid model, to employ "additive method" is better to describe changes of ionosperhic delay in space [1]. In the implementation, vertical ionospheric delay I_v at ionospheric pierce point where the signal path from satellite to receiver crosses the ionosprhic shell, is obtained by GPS dual-frequency measurement data. Data from several reference stations in the area shall be proceeded to acquire I_v of all pierce points. For measurement in the same epoch, establish an array (GCD, $\triangle I_v$), which is made up of Great Circle Distance and difference $\triangle I_v$ ($\triangle I_v = |I_{vi} - I_{vj}|$) of I_v of any two pierce points.

The top graph of Fig. 14.2 is a two-dimensional histogram made by using (GCD, $\triangle I_v$) mentioned above, of which horizontal axis shows GCD and vertical axis shows ionospheric delay difference. Different color used in each bin means different amount of array (GCD, $\triangle I_v$) in this bin. Color choice is confirmed according to the amount of array in the bin. This picture could clearly show changes of ionospheric delay difference varying with GCD.

Based on this, probability distribution curve, just as shown in middle graph in Fig. 14.2, of ionospheric delay difference with different distance could be mapped. In order to read this chart, first choose the curve corresponding to the GCD separation bin you wish to examine, and then select the vertical TEC difference on the horizontal axis that you would like to bound. The intersection of that $\triangle I_v$ value and the chosen curve tells you the probability that a pair of measurements separated by that GCD.

In order to analyze regularity of ionospheric delay difference further, variance changes of ionospheric delay difference, as for different probability, could be mapped, which is shown in bottom graph of Fig. 14.2. In the picture, the value of $\triangle I_v$ for 68 % (1σ), 95 % (2σ), 99 % (2.57σ), 99.9 % (3.29σ) will be calculated respectively. Variance value of different probability, acquired after normalization, varies with the GCD. It could be expected that, if $\triangle I_v$ change is in accord with normal distribution, the four curves shall be overlapped.

14.3.2 Establishment and Analysis of Ionospheric 0th Order Spatial Correlation Model

By using the method said above, GPS dual-frequency data, as for different solar activities and different ionospheric conditions, have been analyzed. The picture in the left side of Fig. 14.2 shows the result of data procession on April 16th 2005. It is found out that:

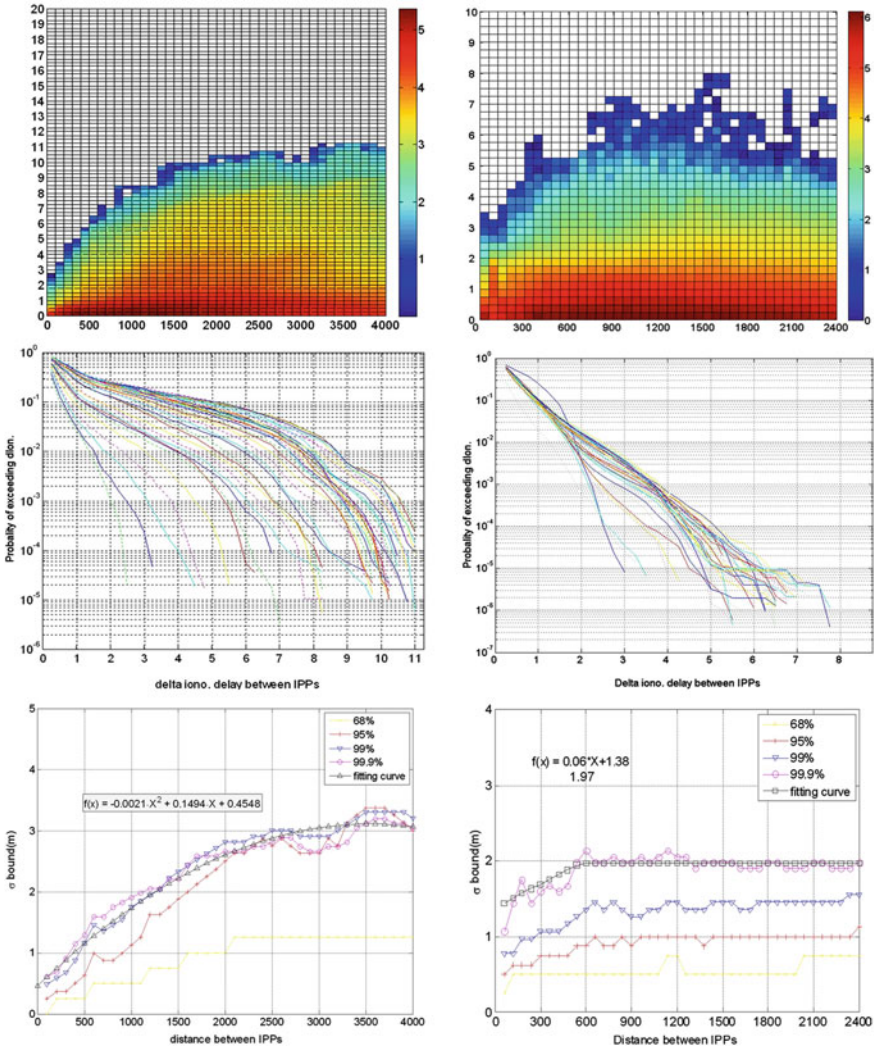


Fig. 14.2 Ionospheric spatial correlation models in China area (*Left* 0th order spatial correlation model, *Right* 1st order spatial correlation model, *Up* relation between ionospheric delay differences (or residuals) and GCD, *Middle* probability curves of ionospheric delay differences (or residuals) of different GCD, *Down* curves of variance of ionospheric delay differences (or residuals) of accumulative probability varying with GCD)

1. When the distance between pierce points is shorter, ionospheric delay still has a larger difference (see up left picture of Fig. 14.2). The difference varies each year for high and low solar activity, and is generally $\sim$3 m in the quiet condition, which indicates the influence of ionospheric anomaly in low-latitude area on ionospheric spatial correlation. When the distance is close, two pierce

points shall have similar ionospheric delay, however, because of existence of ionospheric anomaly in low-latitude area of China, receiver-satellite signal path, which passes the same pierce point from different azimuth and elevation, is most likely to have very different slant ionospheric delay. After this difference transforms into vertical ionospheric delay at the pierce point, it makes pierce points not far away still have a great ionospheric delay difference.

2. From the left picture in the middle of Fig. 14.2, it is found out that probability curves of ionospheric delay difference with different GCD cross, and clearly they show increase to the right-most. It indicates that, because of existence of ionospheric anomaly, even though the distance of two close pierce points, such as in north–south direction, is shorter, gradient change caused by ionospheric anomaly, could also result in a larger ionospheric delay difference. However, two pierce points far away from each other, such as in east–west direction, could also be possible to have a smaller ionospheric delay difference.
3. It is shown in the down left picture of Fig. 14.2 that, for the low-latitude area in China, after the threshold of ionospheric delay differences with different probability is normalized, σ values don't overlap, and in other words, they are not in line with normal distribution. With function fitting analysis for normalized threshold of ionospheric delay differences, it is discovered that, quadratic polynomial function could well describe the regularity. However, in the area of America covered by WAAS system, normalized threshold of ionospheric delay differences is in accordance with normal distribution, and it could well be fitted with linear function [1, 2].

14.4 Ionospheric 1st Order Spatial Correlation Model

14.4.1 Establishing Method of Ionospheric 1st Order Spatial Correlation Model

On the basis of ionospheric 0th order spatial correlation model, ionospheric 1st order spatial correlation model will be studied further.

GPS dual-frequency measurement data is used to extract the vertical ionospheric delay I_v at the ionospheric pierce point. Data from several reference stations within the area will be dealt to acquire I_v for all pierce points. As for as each pierce point is concerned, in the light of planar fitting method, ionospheric delay around is used to obtain ionospheric delay estimate at that pierce point. With the measured ionospheric delay and the estimated ionospheric delay at the pierce point, spatial distribution $\delta vTEC$ of ionospheric delay residual after planar fitting could be acquired.

According to the establishing method of ionospheric 0th order spatial correlation model, ionospheric 1st order spatial correlation model could be established on

the basis of $\delta vTEC$, that is, the function, $\nabla\Delta vTEC = |\delta vTEC_1 - \delta vTEC_2|$, varying in space is established.

14.4.2 Establishment and Analysis of Ionospheric 1st Order Spatial Correlation Model

With the method mentioned above, GPS dual-frequency data are dealt for different solar activities and different ionospheric conditions. The right picture of Fig. 14.2 gives data procession result on September 10th 2005 (the 253 day of year). It could be found out that:

1. It is discovered from the right up picture of Fig. 14.2 that, ionospheric delay residuals, corrected by ionospheric shell model, is about 3–7 m in the year of moderate solar activity. Analysis of data in other period indicates that, the residual is 2–4 m for the year of low solar activity, while it will increase but less than 16 m in the year of high solar activity, besides, if ionospheric disturbance happens, it will continue to increase but less than 18 m.
2. It could be discovered from the right picture in the middle of Fig. 14.2 that, in the year of moderate solar activity such as 2005, probability curves of ionospheric delay residual in the shorter distance show the increase to the right-most. And analysis of data in other time period obviously shows, in the year of high solar activity such as 2003, further increase to the right-most. While in the year of low solar activity, probability curves of ionospheric delay residuals with different distance are clearer, and decrease rapidly. All of these show the effect of ionospheric anomaly in the low latitude, and also proves the assumption of ionospheric shell model has systematic errors. This effect could obviously appear in the year of high solar activity. It also indicates that, in the year of more than moderate solar activity, threshold of ionospheric errors shall increase so as to hold effectively ionospheric delay residuals within the area to meet integrity of the system.
3. It is discovered from the down right picture of Fig. 14.2 that, for the low-latitude area of China, when the threshold of ionospheric delay residuals with different probability is normalized, σ values don't overlap and are not constant values. Function fitting analysis for normalized threshold of ionospheric delay residuals shows that, by means of piecewise function, normalized threshold of ionospheric delay residuals could be well described. And this piecewise function within 600 km in distance is in line with linear variation, while more than 600 km, constant could be used to describe it. This conclusion is different from the result of WAAS system, and the research result in Stanford university shows ionospheric 1st order spatial correlation model, i.e., after planar fitting for ionospheric delay residuals, is a constant [2, 3].

14.5 Conclusion

Ionospheric correlation models are needed for ionospheric grid-based regional augmentation system development, its high-precision ionospheric correction and rigorous integrity performance requirements. Ionospheric spatial correlation models consist of 0th order spatial correlation model and 1st order spatial correlation model, and they respectively describe ionospheric delay differences and ionospheric delay residuals which vary with the distance after planar fitting.

0th and 1st order ionospheric correlation models are established with GNSS data for China area, and analyzed for various ionospheric conditions. The result shows that, in the middle and low latitude of China area, 0th order correlation function (ionospheric delay differences vary with the distance) of ionospheric delay vary with the second-order term of distance, and 1st order correlation function (ionospheric delay residuals after correction of ionospheric grid-based model) is in accord with piecewise function, which is in line with linear variation within 600 km in distance and could be described with a constant beyond 600 km. They are different from the ionospheric spatial correlation models of current regional augmentation system such as WAAS, which is established on the basis of ionosphere change characteristics of the middle latitude. The regional ionospheric characteristics should be taken into account for GNSS integrity application in China area, and the implementation technology should also be designed.

Acknowledgments This work is sponsored by the International Technological Cooperation Projects (No.2011DFA22270) and the National High Technology Research and Development Program (863 Program) of China (No.2011AA120503).

References

1. Hansen A, Peterson E, Walter T, Enge P (2000) Correlation structure of ionospheric estimation and correction for WAAS. In: ION NTM 2000, pp 454–463
2. Hansen A, Walter T, Blanch J, Enge P (2000) Ionospheric spatial and temporal correlation analysis for WAAS: quiet and stormy. In: ION GPS 2000, pp 634–642
3. Walter T, Hansen A, Blanch J, Enge P, et al (2000) Robust detection of ionospheric irregularities. In: ION GPS 2000, pp 209–218
4. Blanch J, Walter T, Enge P (2002) Ionospheric threat model methodology for WAAS. Navigation 49(2):103–107
5. Altshuler ES, Fries RM, Sparks L (2001) The WAAS ionospheric spatial threat model. In: ION GPS 2001, pp 2463–2467
6. Liu D, Zhen W (2012) The effects of China regional ionosphere on satellite augmentation system. Chin J Radio Sci 27(1):195–203

Chapter 15
Study on Pseudolite System for BeiDou Based on Dynamic and Independent Aircrafts Configuration

Guohua Kang, Longyu Tan, Bing Hua and Fengying Zheng

Abstract This paper presents a pseudolite system for BeiDou based on dynamic and independent aircrafts configuration. Firstly, high-altitude and long-endurance unmanned aerial vehicles and airships are needed to obtain the best configuration by analyzing GDOP. Then aiming at the pseudolites' positioning, a new integrated scheme was proposed, where distance between two aircrafts is utilized to improve the positioning accuracy. Simulation results show that GDOP is less than 3.5 and position precision of the pseudolite is 1.2 m in the users' area. The pseudolite system is able to provide moderate navigation service, when BeiDou position system is unavailable to users.

Keywords Pseudolite · BeiDou position system · Dynamic configuration · Integrated position

15.1 Introduction

Satellite navigation with the advantages of all-weather, high-precision, wide coverage [1], has become the main direction for development of navigation technology. But there are still a series of inherent shortcomings: due to the long distance between the satellite and the user, user's signal is weak, so the signal is susceptible to interference or shielding, which may lead to the interruptions of the service [2]; in order to provide WAN services, the signal of the satellite has to

G. Kang (✉) · B. Hua · F. Zheng
Academy of Frontier Science, NanJing University of Aeronautics and Astronautics, NanJing, China
e-mail: kanggh@nuaa.edu.cn

L. Tan
College of Automation Engineering, NanJing University of Aeronautics and Astronautics, NanJing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_15,

cover a wide range, and the satellite itself is in constant movement, therefore, satellite navigation lack the ability to adapt to the complex terrain and cover continuously, which leads to the difficulty in receiving signals in certain terrain conditions; when used in combat zones, due to enemy malicious interference, satellite navigation will face greater difficulties. China's second-generation Beidou satellite navigation system will achieve global navigation in 2020 [3], and it will also face the above problems. Pseudolite augmentation system was first used to augment GPS, now can not only enhance the satellite navigation system to achieve high-precision navigation [4, 5], but also can be independent network within the region to replace satellite navigation system.

This paper presents a kind of pseudolite augmentation system for BeiDou based on dynamic and independent aircrafts configuration of the high-altitude long-endurance UAVs and airships, to enhance area navigation. Concerning the enhancement of space-based navigation network, how to determine the network configuration and improve the positioning accuracy of network nodes are two key issues needing to be addressed first. Firstly, this paper uses high-altitude, long-endurance unmanned aerial vehicle group and airships to build dimensional independent dynamic network; then put forward the method of improving positioning accuracy by the fleet's inner-ranging to help the pseudolite position by itself; finally, present instructions that the constructed pseudolite system can replace the Beidou navigation system as user navigates in the ground area for a short time where the Beidou navigation system fails completely.

15.2 The Independent Dynamic Network of an Airborne Pseudolite

15.2.1 The Line-of-Sight Range of Airborne Pseudolite

In order to expand the coverage of the ground, this paper takes the high-altitude, long-endurance UAV airship as a pseudolite carrier. High-altitude, long-endurance UAV is the unmanned aircrafts of which the flight altitude is above 18 km, the flight time is no less than 24 h, take Global Hawk as an example, its cruising altitude is up to 19.85 km, and airship height and air residence time is even longer. For the Ground-user receiver, there is a minimum elevation angle α, so terrestrial users can only receive a smaller satellite signal range than the high-altitude UAV covers, as shown in the following figure. Simplifying the earth to a ball, taking the semi-major axis of the WGS-84 ellipsoid as the radius R, O as geocentric, S as UAV's sub-satellite point, h as the UAV's height, the corresponding arcs of angle β as the coverage radius of UAV, the corresponding arcs of angle β' as visual user's coverage radius r.

From Fig. 15.1, the radius of the line-of-sight range can be calculated in the following formula:

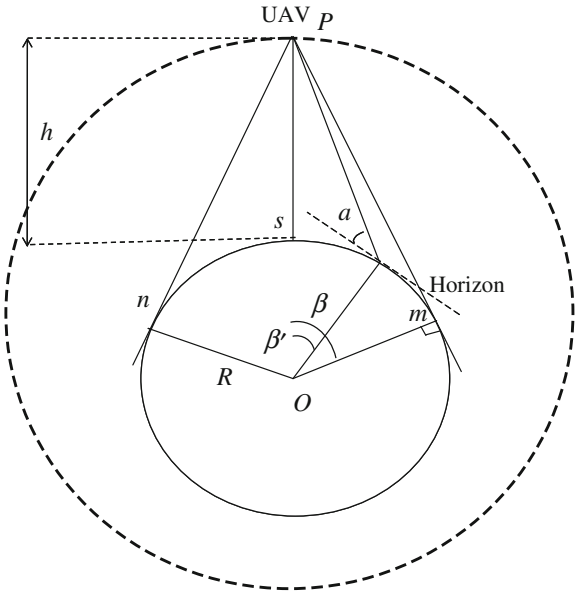


Fig. 15.1 Visual range of pseudolite

$$r = R \times \left[\arccos\left(\frac{R \times \cos(\alpha)}{R+h}\right) - \alpha\right] \tag{15.1}$$

Assume that the height of the UAV from the ground, h =20 km, take the elevation angle $\alpha = 6°$, from formula (15.1), the radius of the line-of-sight range: $r = 168.5$ km.

15.2.2 Airborne Pseudolite Static Networking Configuration

About static configuration of airborne pseudolite networking, there are two constraints mainly needed to be considered:

1. Ground coverage (the intersection of the ground line-of-sight range of the pseudolite) needs to be as large as possible;
2. GDOP needs to be as small as possible.

In fact the two conditions above are mutually contradictory: in order to obtain the GDOP as small as possible, the distance between the point under pseudolite and the user should be as long as possible, but the height of the pseudolite is restricted, which results in the smaller coverage on the ground.

Therefore, the optimal network configuration should base on the actual needs of the area covered and pseudolite's altitude to determine.

Based on the above consideration, take the line-of-sight range of a pseudolite whose height is 20 km as the minimum coverage units, the layout steps to design pseudolites' layout:

1. Set the minimum square $O_1O_2O_3O_4$ around the coverage area A;
2. Take O_1 as sub-satellite point, arrange a pseudolite at the height of h, the radius r of line-of-sight coverage is O_1O_4;
3. Do as step 2, respectively, take O_2, O_3, O_4 as the sub-satellite point, arrange a pseudolite at the height of h above O_2, O_3, O_4
4. Take the center O of the square as the sub-satellite point, arrange a pseudolite at the height of $h'(>h)$ above O. The line-of-sight coverage radius of the pseudolite is $r'(>r)$.

There are following advantages if using the layout above:

1. Ensuring the full coverage of the area. As can be seen from Fig. 15.2, the range covered by the pseudolite is $G_1G_2G_3G_4$, including a region A, and there is spare range;
2. The configuration of 5 pseudolites to make sure that the consisted pseudolite system has an independent navigation capability, and has a margin;
3. The aerial configuration of 5 pseudolite is a 5 icosahedral with a good GDOP basis.

15.2.3 Performance Analysis of Independent Dynamic Network

With the pseudo-static networking configuration determined, this paper adopted dynamically layout to build Beidou satellite system: 4 pseudolite above the sub-satellite point of $O_1O_2O_3O_4$, using high-altitude, long-endurance unmanned aircraft maintaining spacing while doing variable speed circular motion in the same circumference, the circumference is 20 km above the ground, a diameter of about 200 km, the center sub-satellite point is O; the pseudolite platform above O using high-altitude airship, relative still to the point O, 30 km away from the ground. For

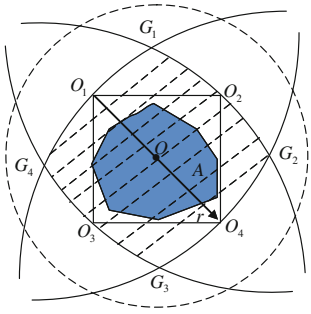


Fig. 15.2 Covering range of pseudolite satellites

easier description, the airship and high-altitude, long-endurance UAV are unified with the "UAV" to refer to below.

Due to the distance between pseudolites, and distance between pseudolite and the user are both very short (10–300 km), it can be assumed the ground plane to plane, calculated GDOP using the local origin of the Cartesian coordinate system. This can simplify the calculation, meet more needs of practical engineering [6]. GDOP is calculated as follows:

$$GDOP = \sqrt{tr(\mathrm{G}^T\mathrm{G})^{-1}} = \sqrt{q_{11} + q_{22} + q_{33} + q_{44}} \tag{15.2}$$

In the above formula, G is the position matrix array, $q_{ii}(i = 1, \ldots 4)$ is the diagonal elements [7, 8].

Take the 200 × 200 km ground area under the pseudolite system as the users' area, through the Matlab simulation, the 3D map of GDOP value changes of the area shown in Fig. 15.3, the corresponding contour map shown in Fig. 15.4.

As can be seen from the perspective view of Fig. 15.3, the nearer to the center region of the user, the smaller is the GDOP value, the better is the navigation performance; in the contour map Fig. 15.4, the rhombic area where PDOP is less than 3.5, is the actual usable area of the users' common-view satellites which is corresponding to $G_1G_2G_3G_4$ in Fig. 15.3. In fact, due to the dynamic rotation of pseudolites in the sky, Figs. 15.3, 15.4 only refer to the GDOP at a certain moment, that is to say, the PDOP of the user changes every moment.

15.3 The High-Precision Positioning of Airborne Pseudolite

With the UAV group configuration determined, the airborne pseudolite system positioning accuracy depends on the accuracy of pseudolite itself, so the high-precision positioning of UAV becomes the key to the system. The technology of

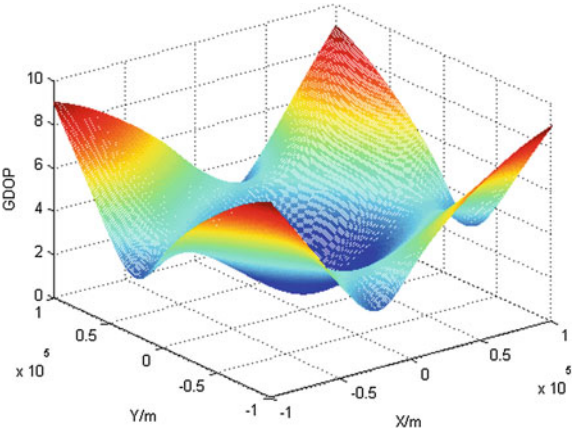


Fig. 15.3 Three-dimension graph of GDOP

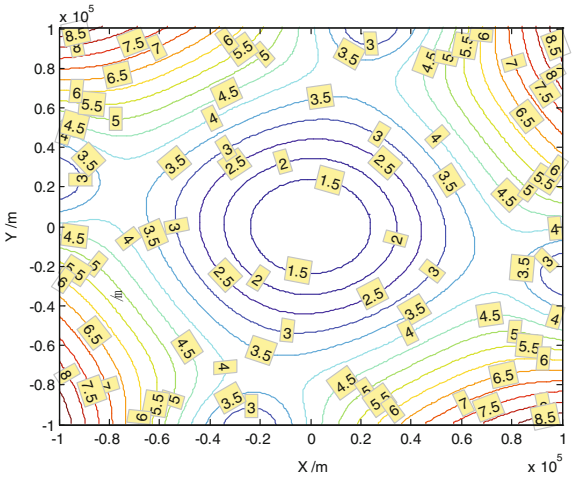


Fig. 15.4 Contour map of GDOP

navigation system carried by UAV's own platform is mature, usually using the combined navigation technology [8] of the inertia/satellite navigation; as a result of using the fleet Networking, there is rich information can be taken advantage of. Due to this, this paper proposes a combination of positioning method based on fleet inner-ranging to improve the positioning accuracy, that is, the distance between the UAVs would be used as observations, and introduced to combo Positioning System to fuse with the information, in order to further improve the positioning accuracy of the fleet.

The state of the system equation:

$$X(t) = F(t)X(t) + G(t)W(t) \tag{15.3}$$

Using 20 Victoria state variables, respectively, as error angle of platform, speed error, position error, the random constant of gyro, first-order Markov process noise of gyro, first-order Markov process noise of accelerometer, distance corresponding to equivalent clock error, distance rate corresponding to equivalent clock frequency error. The relevant state matrix $F(t)$, the system noise matrix and its discretization can turn to Ref. [3].

15.3.1 The Filter Model of INS/Fleet Inner-Ranging

Take two UAVs as an example, assuming one is a measured UAV, the other UAV is as reference. The distance between the two is the function of their positions. In order to decoupling with the referenced UAV when estimating the state of the measured UAV, the referenced UAV not only send its own state to the measured UAV, as well as the accuracy evaluation of the state, that is the error covariance array. According to error covariance matrix of the referenced UAV, the distance

between the fleet can be used, and the state of the referenced UAV's state is no longer estimated in the filtering process. In order to reduce the amount of calculation, as well as reducing the interdependence among the UAVs, the distributed sequential processing method is used. "Distribution" means that every UAV does navigation solver on its own according to the ranging value between the fleet and status information sent by other UAVs. "Order" means that measured value within the same measurement cycle would be respectively input to the Kalman filter to fuse with information. The specific process is shown in Fig. 15.5.

Take two UAVs as example (one denoted as UAV a, another as UAV b), the filtering process is as follows:

The distance $\Delta\rho_{ab}$ between fleet can be expressed as:

$$\Delta\rho_{ab} = \mathrm{H}_a X_a + \mathrm{H}_b X_b + \xi \tag{15.4}$$

Assuming UAV b as the referenced UAV, while doing filtering estimation of the UAV a taking the position error of UAV b as noise, the equation above can be further expressed as follows:

$$\Delta\rho_{ab} = \mathrm{H}_a X_a + \xi(\mathrm{H}_b X_b) + \xi \tag{15.5}$$

$\xi(\mathrm{H}_b X_b)$, denotes noise processing, the specific implementation can be seen in the following filtering equation group.

The filter equations for INS/Fleet ranging, is shown in formula (15.11). $H_{a,k}$ is the measurement matrix for the estimated UAV, $P_{a,k/k-1}$ is the state covariance matrix for the estimated UAV, $H_{\mathrm{b},k}$ is measurement matrix for the referenced UAV, $P_{\mathrm{b},k/k-1}$ is the state covariance matrix, $\hat{X}_{k/k-1}$ is a one-step prediction, $\mathrm{F}_{k,k-1}$ is the system matrix, $G_{k,k-1}$ is the system noise array, Q_{k-1} is the covariance

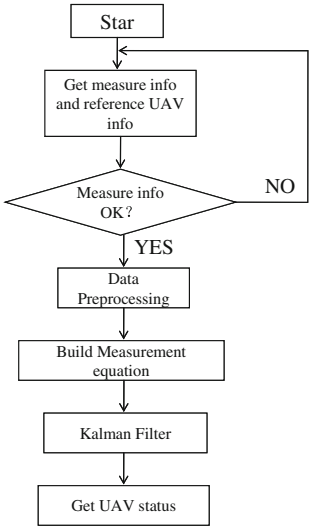


Fig. 15.5 Distribution and order filter process

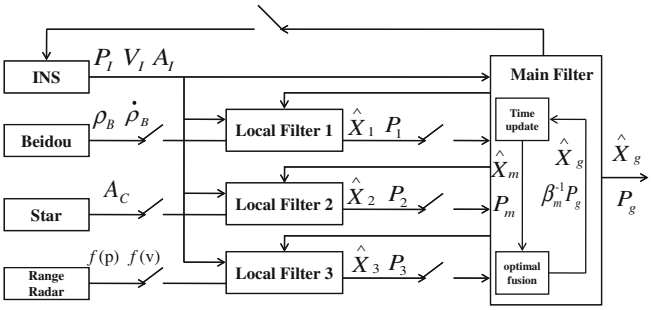


Fig. 15.6 Integrated position scheme

matrix for estimated error, R_k is the covariance matrix for the measurement noise, Z_k is observables.

$$\begin{cases} \hat{X}_{k/k-1} = \mathrm{F}_{k,k-1}\hat{X}_{k-1/k-1} \\ P_{k/k-1} = \mathrm{F}_{k,k-1}P_{k-1/k-1}\mathrm{F}_{k,k-1}^T + G_{k,k-1}Q_{k-1}G_{k,k-1}^T \\ K_k = P_{k/k-1}H_{a,k}^T[H_{a,k}P_{a,k/k-1}H_{a,k}^T + R_k + H_{\mathrm{b},k}P_{\mathrm{b},k/k-1}H_{\mathrm{b},k}^T]^{-1} \\ \hat{X}_{k/k} = \hat{X}_{k/k-1} + K_kZ_k \\ P_{k/k} = (\mathrm{I} - \mathrm{K}_kH_{a,k})P_{a,k/k-1}(\mathrm{I} - \mathrm{K}_kH_{a,k})^T + K_k(R_k + H_{\mathrm{b},k}P_{\mathrm{b},k/k-1}H_{\mathrm{b},k}^T)K_k^T \end{cases} \tag{15.6}$$

Similarly, to the inter-distance observation, UAV *a* as the referenced UAV, while UAV *b* as the estimated, similar amendment to UAV *b*, the specific filtering process is just as the above.

15.3.2 The Enhanced Combined Positioning Scheme Based on Fleet Ranging

A positioning program that every UAV is based on fleet ranging to improve the positioning accuracy of Integrated Positioning, that is using the federal filtering algorithm, taking inertial navigation system as a reference system, introducing the output pseudorange and pseudorange rate of Beidou satellite navigation system to the local filter 1, the introduction of the output attitude information of the star sensor to a local filter 2, introducing the distance between the fleet obtained by radio ranging information to local filter 3, using fusion—feedback mode [9], the specific implementation is shown in Fig. 15.6. P, V, A are respectively UAV's position, velocity and attitude, ρ_B, $\dot{\rho}_B$ are respectively output pseudorange and pseudorange rate from the Beidou navigation systems, A_C stands for output attitude information from star sensor, $f(\mathrm{P})$ is the function of position.

15.3.3 Local Filtering Measurement Equation

The local filter 1 is based on integrated navigation of INS/Beidou pseudorange and pseudorange rate, and its specific derivation process for measurement equations can turn to Ref. [10, 11], not repeating here.

15.3.3.1 The Measurement Equation for Local Filter 2

Local filter 2 is a filter based on integrated navigation of INS/star sensor, the measurement equation:

$$Z_2 = H_2 X + V_2 \tag{15.7}$$

Express platform error angles of the Eastward, the Northward, and the Universe as $[\varphi_E \quad \varphi_N \quad \varphi_U]^T$, and rolling, pitching, heading attitude error angles are denoted by $[\delta\gamma \quad \delta\theta \quad \delta\psi]^T$, the conversion from platform error angle to the attitude error angle can be expressed as following relations [12]:

$$[\delta\gamma \quad \delta\theta \quad \delta\psi]^T = H_\varphi [\varphi_E \quad \varphi_N \quad \varphi_U]^T \tag{15.8}$$

In the above formula, $H_\varphi = \frac{-1}{\cos\theta}\begin{bmatrix} \sin\psi & \cos\psi & 0 \\ \cos\psi\cos\theta & -\sin\psi\cos\theta & 0 \\ \sin\psi\sin\theta & \cos\psi\sin\theta & -\cos\theta \end{bmatrix}$

So the measurement matrix is:

$$H_2 = [H_\varphi \quad 0_{3\times 17}] \tag{15.9}$$

Attitude angles obtained by Inertial Navigation system and the Star-sensitive are respectively denoted by $[\gamma_I \quad \theta_I \quad \psi_I]^T$ and $[\gamma_G \quad \theta_G \quad \psi_G]^T$, so:

$$Z_2 = [\gamma_I - \gamma_G \quad \theta_I - \theta_G \quad \psi_I - \psi_G]^T \tag{15.10}$$

15.3.3.2 Measurement Equation of Local Filter 3

Local filter 3 is the filter based on integrated navigation of INS/Fleet inter-ranging. Measurement equation of local filter 3:

$$Z_3 = H_3 X + V_3 \tag{15.11}$$

Range only once between any two UAVs, and 10 ranging values are obtained, still take UAV a and UAV b as an example to get the measurement equation.

Set (x_a, y_a, z_a), (x_b, y_b, z_b) as the solid Coordinates for UAV a and UAV b, then distance between them is:

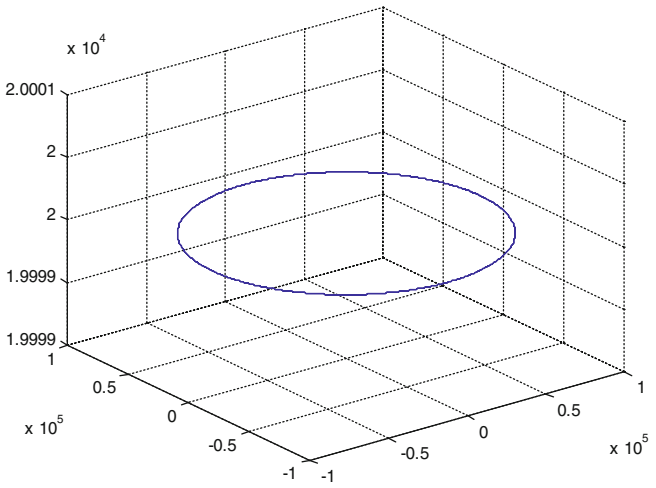


Fig. 15.7 Flight path of UAV

Table 15.1 Simulation parameters

Simulation parameter	Values
Gyro white noise	0.01 °/h
The gyro first-order Markov noise	0.01 °/h
The gyro first-order Markov time constant	3,600 s
Accelerometer white noise	10^{-4} g
Accelerometer first-order Markov noise	10^{-4} g
Accelerometer time constant of a first-order Markov	1,800 s
The white noise rms of Beidou pseudorange	5 m
The white noise of Beidou pseudorange rate	0.01 m/s
Star sensor's accuracy	$5''$
Radio's ranging accuracy	1 m
Simulation time	1,800 s

$$\Delta\rho_{ab} = \sqrt{(\mathrm{x}_a - \mathrm{x}_b)^2 + (\mathrm{y}_a - \mathrm{y}_b)^2 + (\mathrm{z}_a - \mathrm{z}_b)^2} \tag{15.12}$$

Thus, the measurement matrix:

$$H_3 = \left[0_{1\times 6} \quad \frac{\partial \Delta\rho_{ab}}{\partial x_a} \quad \frac{\partial \Delta\rho_{ab}}{\partial y_a} \quad \frac{\partial \Delta\rho_{ab}}{\partial z_a} \quad 0_{1\times 11} \right] \times C_n^e,$$

in the quota, C_n^e is the transfer matrix from the navigation coordinate system to the coordinate system of the Earth's solid company.

$$Z_3 = [\Delta\rho_I - \Delta\rho_c] \tag{15.13}$$

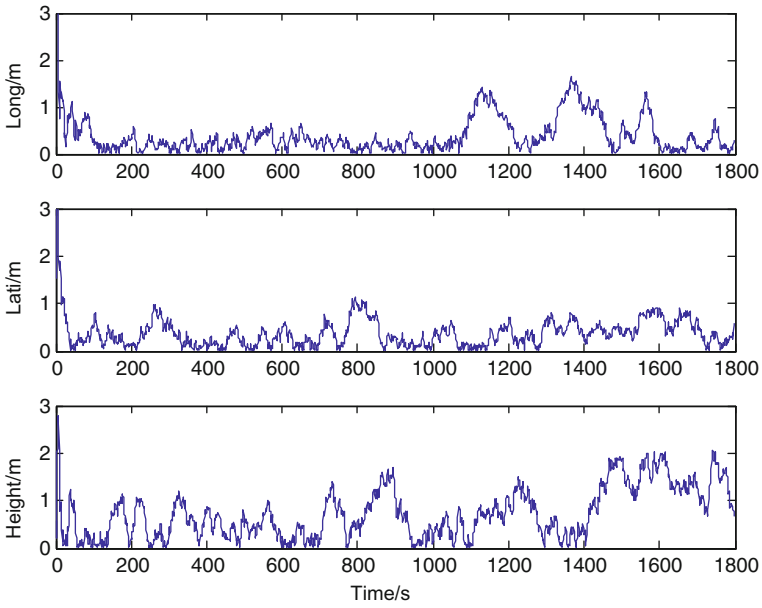


Fig. 15.8 Position error

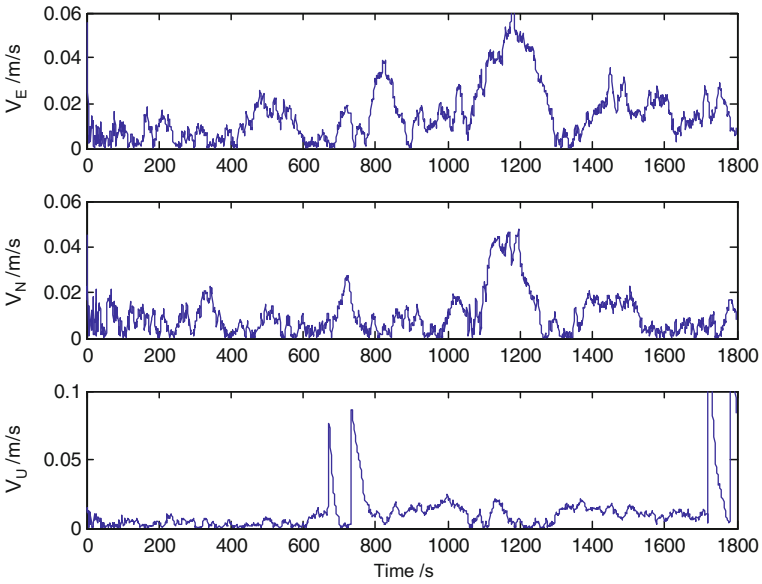


Fig. 15.9 Velocity error

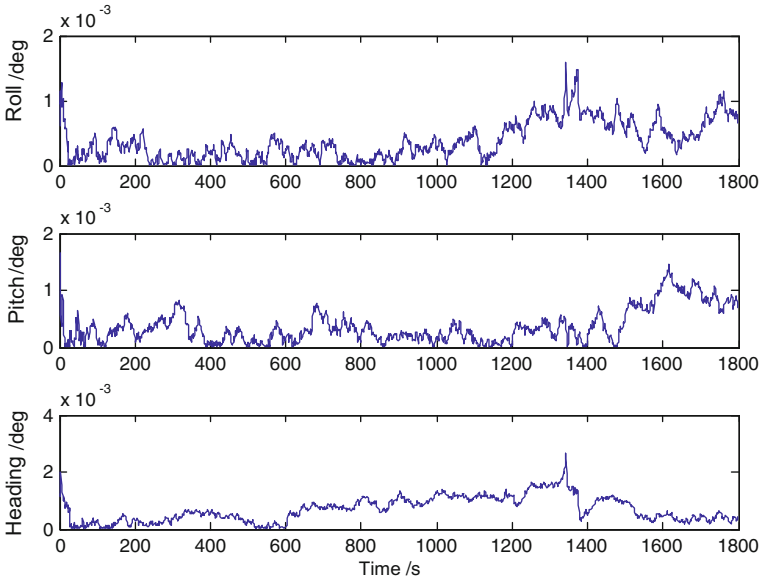


Fig. 15.10 Attitude error

In the quota, $\Delta\rho_I$ is the distance between the two UAVs obtained from inertial navigation system solver, while $\Delta\rho_c$ is the value for distance between the two UAVs measured by the radio.

15.3.4 Simulation Analysis

In order to verify the effectiveness of the method for integrated navigation by using fleet inter-ranging to improve the positioning accuracy, a complete set of digital simulation platform is established, including the flight path simulator, every sensor simulation module, Federated Filter and so on. Carry out simulation experiments of high-precision positioning on airborne pseudolite, then analyze the effectiveness.

15.3.4.1 Setting the Track

According to the above context, set the track as follows: the airships are still, and the four high-altitude, long-endurance unmanned aerial vehicle interval the same in the same circumference, for variable speed circular motion with the same law of motion, including acceleration, cornering, level flight maneuvers. Among them, one UAV's track is shown in Fig. 15.7.

15.3.4.2 Simulation for INS/Beidou/Star/Fleet Inner-Ranging Integrated Navigation

The simulation parameters are set as shown in Table 15.1.

As space is limited, only the curves for position error, speed error and attitude error of one UAV, are shown in Figs. 15.8, 15.9 and 15.10; from the statistics obtained through the error data, the position error is 1.20 m, the speed error is 0.032 m/s. The rest UAVs' error accuracy and the above UAV is basically the same.

15.4 Conclusion

In this paper, a pseudolite system for BeiDou based on dynamic and independent aircrafts is discussed, including the dynamic network formation for pseudolite and the method for integrated navigation by using fleet inter-ranging to improve the positioning accuracy. And fleet inter-ranging filter model is estabished, the observation equations for every local filter of the integrated positioning are derived, and detailed simulations have been performed. The results of simulation indicate that the constructed pseudolite system is able to provide moderate navigation service instead of BeiDou, when BeiDou positioning system is unavailable to users.

References

1. Li L (2009) Research of the deep integration of GPS/SINS navigation system. Shanghai Jiao Tong University, Shanghai
2. Malleswaran M, Angel Deborah S, Manjula S (2010) Integration of INS and GPS using radial basisfunction neural networks for vehicular navigation. In: 11th international conference control, automation, robotics and vision 010, pp 2427–2430
3. Liu J, Zeng Q, Zhao W (2010) Theory and application of navigation system. Northwest Industrial University Press, Xi'an
4. Cobb HS (1997) GPS pseudolites: theory, design, applications. USA:Standford University, California
5. Lee HK, Wang JL, Rizos C, et al (2002) Analysis of pseudolite augmentation for GPS airborne applications. In: IONGPS, Portland, Oregon, American
6. Yue Y, Lie H (2009) GAO guan-gen. Precision analysis of pseudolite positioning system. Chin J Sens Actuators 22(10):1432–1436
7. Meng J (2007) Study on pseudolite position technology and configuration. Zhengzhou:Information Engineering University
8. Yarlagadda R, Ali L, Al-Dhahir N, Hershey J (2000) GPS GDOP metric. IEE Proc Radar Sonar Navig 147(5):259–264
9. Zhnag N (2006) Altitude long-endurance UAV integrated navigation system. Northwestern Polytechnical University, Xi'an

10. Yuan G, Yuan K, Zhang H (2011) A variable proportion adaptive federal kalman filter for INS/ESGM/GPS/DVL integrated navigation system. In: 4th international joint conference on computational sciences and optimization 2011, pp 976–981
11. Yuan J, Fan S, Liu J (2011) Research on navigation performance of GPS/INS tightly integrated system. Aerosp Electron Warfare 27(2):17–21
12. Zhou J, Yuan J, Luo J (2009) Fully intergrated SINS/GPS navigation system based on federated filters. J Syst Simul 21(6):1562–1569

Chapter 16
GNSS Satellite Selection Algorithm Revisited: A Weighted Way with Integrity Consideration

Liang Li, Hong Yuan, Chao Yuan, Dongyan Wei and Wei Liu

Abstract With the promising future of Global Navigation Satellite System (GNSS), the quantity of available satellites could dramatically increase to more than 35. Tracking all available satellites simultaneously leads to unacceptable computation burden for standalone receivers. Most satellite selection methods were therefore proposed to release the computation load with the minimum satellites set to maintain endurable positioning accuracy under the least Dilution of Precision (DOP) loss criterion. In terms of safety-of-life (SoL) navigation services, the integrity monitoring is an indispensable function. The traditional satellite selection methods should therefore be enhanced to take account of positioning accuracy and the potential integrity loss from some fault satellite, which may originates from satellite clock anomaly drift or the deformation of signals in space, e.g. ionosphere scintillation. Giving the fact that the loss of some satellite may have little difference on the DOP, but could impose a negative influence on the integrity performance, satellite selection should therefore give an insight into the selection criterion and the determination of the threshold values. A novel satellite selection algorithm with using weighting method is presented to select the optimal satellites set with satisfying the requirement in both accuracy and integrity. The selection threshold values for DOP and integrity is deduced from the required navigation performance (RNP) defined by International Civil Aviation Organization (ICAO). The performance of the proposed method was validated with using the simulated GNSS constellations. The accuracy and integrity also were compared with the typical existing satellite selection methods, i.e., maximum volume algorithm and quasi-optimal satellite selection algorithm, etc. The simulation results demonstrate that weighted satellites selection method could provide optimal satellites set with not only the better accuracy and less integrity loss, but also with a reasonable computation burden, and thus could be an edge option for the GNSS satellite selection methods.

L. Li (✉) · H. Yuan · C. Yuan · D. Wei · W. Liu
Academy of Opto-electronics, Chinese Academy of Science, Beijing 100094, China
e-mail: lilianghrb@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_16,

Keywords GNSS · Satellite selection · Accuracy · Integrity · DOP

16.1 Introduction

With the promising future of Global Navigation Satellite System (GNSS), i.e., the integration of US GPS, Russian GLONASS, Chinese Beidou Navigation Satellite System (BNSS) and European Galileo, the quantity of orbiting satellites could dramatically increase to more than 100, it can therefore be anticipated that the number of available satellites can be much more than ever even if there are frequent blockages. It is no doubt that GNSS could be a prominent technology to provide higher accuracy and more reliable positioning solution than standalone constellation. However, tracking all visible satellites simultaneously leads unacceptable computation burden for standalone receiver and even worse positioning accuracy if some satellite signals contains more severe measurement error than the others. This is one of significant origination for satellite selection. One of the widely utilized satellite selection criterions is the cut-off elevation angle, which is effective in mitigating severe multipath error from low elevation satellite [1]. Nevertheless, the extreme of elevation selection (e.g. the highest elevation satellite selection algorithm) induce poor geometry dilution of precision (DOP), i.e., the implication of amplification of measurement errors. Most satellite selection methods were therefore proposed to minimum DOP loss due to the reduction of the number of satellite [2, 3]. In satellite navigation based positioning concept, four satellites measurements at least are necessary to obtain position solution. One of obvious satellite selection options is therefore to find four satellites with the minimum DOP loss [4], the computation burden however increase as more available satellites with the multiple GNSS constellations. The maximum volume algorithm (MVA) is therefore proposed to select four satellites that minimize the DOP with the reasonable computation load, but is limited to only four satellites because of regular tetrahedron shape and lack of the ability of integrity, which is sensitive to the measurement fault contained in the selected satellites signals that may arise from the deformation of signal-in-space, e.g. ionospheric scintillation, or the anomaly of satellite clock drift. Another typical selection criterion is to eliminate the *redundancy* of satellites [3]. The selection process is based on the calculation of redundancy cost function that determines the geometric closeness of one satellite to the other visible satellites, and eliminates the satellite with maximum redundancy until satisfying the predefined number of satellites. These existing satellite selection methods mostly pay too much attention to the satellite geometry, but little consideration on the effect of losing one satellite on the integrity performance. However, in terms of safety-of-life (SoL) navigation services, the integrity monitoring is an indispensable function [5, 6]. The performance of integrity monitoring is assessed by accuracy, integrity, continuity, and availability (i.e., Required Navigation Performance, RNP) [7, 8], and closely related with satellite geometry [9]. The satellite selection method should therefore take

account of positioning accuracy and integrity requirement. On one hand, giving the fact that the loss of some satellite may have little attribution to the DOP, but could impose a negative influence on the integrity level, satellite selection should therefore give an insight into two issues: Which satellite signals should be tracked simultaneously? How many satellites should be selected? The solutions of these issues correspond to the selection criterion and the selection threshold values, respectively. On the other hand, RNP for civil aviation shows that various navigation phases have different horizontal or vertical performance requirement [8], e.g. en-route, terminal and non-precision approach pay more attention to the horizontal navigation performance, while vertical performance is the primary concern for the precision approach phase (e.g. CAT I/II/III). The RNP difference of various navigation phases defined by ICAO (International Civil Aviation Organization) implies the adjustment of selection criterion.

In order to maintain the availability, the output of integrity monitoring, i.e., protection level (PL) should be within the range of alert limit (AL) to avoid creating hazardous misleading information. Furthermore, the predefined evaluation parameters of integrity performance (e.g. continuity risk and integrity risk) have to be satisfied by comparing the AL and PL without giving the evaluation reference. All these mentioned requirements for PL imply that the satellite selection threshold values could be deduced from the predefined RNP parameters. A novel satellite selection algorithm with using weighting method is therefore proposed to develop the selection criterion with integrity consideration. The selection threshold values of DOP and integrity for the weighting method is deduced according to the predefined RNP parameters.

The rest of the paper is organized as follows. In the first section, the necessity of satellite selection criterion adjustment is revealed with the simulation method. Then, the weighted satellite selection method is proposed and analyzed with the deduction of the threshold values from RNP parameters. The experiments result and conclusion of the proposed method are discussed in the third and forth section.

16.2 The Necessity of Satellite Selection with the Consideration of Integrity

16.2.1 The Necessity of Satellite Selection

The necessity of satellite selection is not quite obvious when the US GPS is the sole full operation capability (FOC) constellation since the computation power is within the capability of general-purpose processors, as the available satellites for standalone receiver are no more than twelve. However, with the rapid development of integration between GPS, GLONASS, Galileo and BNSS, GNSS interoperability is the inevitable tendency and leads to mass of available satellites as high as more than 35 [10]. However, the signal quality of some satellites is more valuable than others, e.g., multipath and atmospheric error from low elevation satellite,

sever thermal noise error induced by low carrier-to-noise (C/N_0), etc. The low signal quality of satellite has a direct negative influence on the range accuracy. One important reason for satellite selection is to keep the high-quality satellite signals utilized in the positioning process. Furthermore, it is no doubt that the positioning accuracy could be enhanced with tracking all the satellite signals in-view under the premise of stable user equivalent range error, but the processing efficiency will be low as more computation power is needed to satisfy the real-time navigation requirement. Another reason for satellite selection is therefore proposed to release the limitation between computation power and positioning accuracy. Although the positioning error could be worse after satellite selection, the accuracy requirement can be still satisfied with deliberately selected satellites set, which is the foundation of satellite selection.

As satellite geometry interacts with the measurement error, and DOP parameters are used to characterize the effect of satellite geometry on positioning accuracy, the mathematical linear pseudorange model is firstly given for the convenient introduction of DOP at next:

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{X} + \epsilon \tag{16.1}$$

In which $\boldsymbol{y}$ is the $n \times 1$ pseudorange measurements, $\boldsymbol{G}$ is the $n \times 4$ geometry matrix of user to satellite, $\boldsymbol{X}$ is 4×1 the unknown user position and clock error, $\boldsymbol{\varepsilon}$ is the $n \times 1$ pseudorange error and follows the normal distribution $N(0, \sigma^2)$. Matrix $\boldsymbol{A}$ can be computed with the following equation to calculate different DOPs.

$$\boldsymbol{A} = \left(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\right)^{-1} \tag{16.2}$$

GDOP, PDOP, HDOP and VDOP are defined as:

$$\left.\begin{aligned} GDOP &= \sqrt{\mathrm{trace}(\boldsymbol{A})}, & PDOP &= \sqrt{\sum_{i=1}^{3}\boldsymbol{A}_{ii}}, \\ HDOP &= \sqrt{\sum_{i=1}^{2}\boldsymbol{A}_{ii}}, & VDOP &= \sqrt{\boldsymbol{A}_{33}} = \sqrt{PDOP^2 - HDOP^2} \end{aligned}\right\} \tag{16.3}$$

In which $\boldsymbol{A}_{ij}$ is the element of ith row and jth column in $\boldsymbol{A}$. As receiver clock error could be different due to the quality of crystal oscillator, this paper focuses on PDOP to study the effect of receiver-satellite geometry on the three dimensional position solutions, which will be commonly used at next.

16.2.2 The Necessity of Integrity Consideration in Satellite Selection

Based on the above analysis, satellite selection is developed to release the computation load and select high quality satellite signals, and minimize positioning

accuracy loss due to the reduction of the quantity of the satellites. As far as SoL navigation services are concerned, e.g. civil aviation and hazardous chemical materials transportation, integrity is of highly necessary and an indicator for the quality of satellite signal. Since integrity is directly affected by the satellite geometry, satellite selection should therefore take account for integrity loss.

With the linear measurement model, positioning solution and the Protection Level (PL) can be obtained. In terms of standalone receiver, the *xPL* (*x* could be H for *Horizontal* or V for *Vertical*) can be calculated as follows [11]:

$$xPL = xSlope_{\max}\sigma\sqrt{\lambda} \tag{16.4}$$

where λ is the non-central parameter of the non-central Chi square distribution under the single-fault hypothesis. Subscript max means the maximum operation. *xSlope* is the geometry parameter and it can be directly computed from one specific satellite user geometry, according to the following equations in horizontal and vertical plane respectively:

$$HSlope_i = \sqrt{\frac{\boldsymbol{H}_{1i}^2 + \boldsymbol{H}_{2i}^2}{\boldsymbol{S}_{ii}}} \tag{16.5}$$

$$VSlope_i = \frac{|\boldsymbol{H}_{3i}|}{\sqrt{\boldsymbol{S}_{ii}}} \tag{16.6}$$

where $\boldsymbol{H} = (\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathrm{T}}$, $\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{G}\boldsymbol{H}$. It can be seen from Eq. (16.4) that *xSlope* is the parameter that connect *xPL* with the satellite geometry, as λ are deduced by navigation phase and can be determined off-line. It is noted that the integrity performance is inversely proportional to the increasing of *xPL*, which means the integrity loss will be more severe with the larger *xPL*. The same way as PDOP is used to assess the positioning accuracy due to satellite selection, *xSlope* is the indicator to characterize the integrity.

16.3 The Weighted Algorithm for Satellite Selection

Based on the above analysis and evaluation of the limitation of traditional selection criterions, a weighted satellite selection algorithm is developed to satisfy both the positioning accuracy and integrity requirement.

16.3.1 Threshold Values for PDOP and Integrity

As AL is the threshold values of PL, which means PL should be within the protection of AL to maintain the availability requirement, i.e., $xPL \leq xAL$ have to

be satisfied. Combining Eq. (16.4), the threshold values of *xSlope* can be computed as:

$$Th_{xSlope} = \frac{xAL}{\sigma\sqrt{\lambda}} \tag{16.7}$$

The non-central parameter λ can be derived from the probability of missed detection P_{md} and false alarm P_{fa}, which can be expressed as:

$$P_{fa} = 1 - \int_0^{T_d} f_{\chi^2(n-4)}(t)\mathrm{d}t \tag{16.8}$$

$$P_{md} = \int_0^{T_d} f_{\chi^2(n-4,\lambda)}(t)\mathrm{d}t \tag{16.9}$$

In which $f_{\chi^2(n-4)}(t)$ is the probability of density function (pdf) of central Chi square, and $f_{\chi^2(n-4,\lambda)}(t)$ is the pdf of non-central Chi square. The numerical solution of λ can be obtained from Eq. (16.9) with the threshold value T_d computed from Eq. (16.8). The threshold value for integrity can be obtained from Eqs. (16.7), (16.8) and (16.9).

As is proven in [12] that $xSlope = \Delta xDOP^2$, in which $\Delta xPDOP$ is the PDOP loss due to the removing of some satellite. The threshold value of DOP loss can therefore be equivalently expressed as:

$$Th_{\Delta xDOP} = \sqrt{Th_{Slope}} \tag{16.10}$$

From Eq. (16.3), we have

$$PDOP_i^2 = HDOP_i^2 + VDOP_i^2 \tag{16.11}$$

The first derivative in time of Eq. (16.11) is therefore:

$$PDOP_i \cdot \frac{\mathrm{d}PDOP}{\mathrm{d}t} = HDOP_i \cdot \frac{\mathrm{d}HDOP}{\mathrm{d}t} + VDOP_i \cdot \frac{\mathrm{d}VDOP}{\mathrm{d}t} \tag{16.12}$$

Let

$$\left.\begin{aligned} \frac{\mathrm{d}PDOP}{\mathrm{d}t} &= \Delta PDOP_i = PDOP_i - PDOP_o \\ \frac{\mathrm{d}HDOP}{\mathrm{d}t} &= \Delta HDOP_i = HDOP_i - HDOP_o \\ \frac{\mathrm{d}VDOP}{\mathrm{d}t} &= \Delta VDOP_i = VDOP_i - VDOP_o \end{aligned}\right\} \tag{16.13}$$

In which subscript o and i is the value before and after removing ith satellite. By substituting Eq. (16.13) into Eq. (16.12), the following can be obtained

$$\Delta PDOP_i = \frac{HDOP_i}{PDOP_i}\Delta HDOP_i + \frac{VDOP_i}{PDOP_i}\Delta VDOP_i \tag{16.14}$$

With the threshold value of $\Delta HDOP$ and $\Delta VDOP$, the threshold value of PDOP can be calculated as follows:

$$Th_{\Delta PDOP} = \frac{HDOP_i}{PDOP_i} Th_{\Delta HDOP} + \frac{VDOP_i}{PDOP_i} Th_{\Delta VDOP} \tag{16.15}$$

There are some issues need to be clarified before obtain the specific selection threshold values. Firstly, the threshold values for DOP and integrity relate with RNP parameters, the threshold values vary with different navigation phases. Secondly, as P_{md} and P_{fa} are not directly provided by GNSS Standards and Recommended Practices(SARPs) [13], it is therefore necessary to transform the integrity risk (IR) and continuity risk (CR) requirement to the P_{md} and P_{fa} requirement. With using the calculation method from [14], the specific P_{fa} and P_{md} requirement of modes from En-route to AVP are listed in Table 16.1 as the availability of GNSS based RAIM have difficulty in satisfying the RNP of precision approach [14].

It can be seen from Table 16.1 that en-route, terminal and NPA focus on the horizontal navigation performance, while APV has stricter performance requirement in the vertical direction.

16.3.2 Weighted Satellite Selection Algorithm

Since the satellite geometry related with DOP and integrity, it is therefore necessary to modify the traditional selection criterion with using the threshold values in the previous section. As seen from Table 16.1, the navigation modes from En-route to NPA only have specific horizontal requirement, while APV and precision approach have much stricter requirement at both horizontal and vertical direction. As known from previous section, the satellite selection result with integrity requirement in the horizontal and vertical direction could be different, which is an implication that selection criterion should be alternative at different flight phases.

To take account of PDOP and integrity on the satellite selection, the cost function for removing the ith satellite can be generally expressed as:

Table 16.1 P_{fa} and P_{md} requirements at different flying phases

Phases	ALs	CR	IR	P_{fa}	P_{md}
En-route	3.7 km(H)	10^{-5}/h	10^{-7}/h	3.33×10^{-7}	10^{-3}
Terminal	1.85 km(H)				
NPA	220 m(H)				
APV I	40 m(H)	8×10^{-6}/15 s	2×10^{-7}/150 s	4×10^{-6}	
	50 m(V)				
APV II	40 m(H)				
	20 m(V)				

$$J_i = w_{PDOP} \cdot \frac{\Delta PDOP}{Th_{\Delta PDOP}} + w_{HI} \cdot \frac{HSlope}{Th_{HSlope}} + w_{VI} \cdot \frac{VSlope}{Th_{VSlope}} \tag{16.16}$$

In which the selection weights, i.e., w_{DOP} for PDOP loss $\Delta PDOP$, and w_{HI} and w_{VI} for the slope loss which reflect the integrity loss, satisfy the following requirements:

$$\left.\begin{array}{ll} w_{DOP} + w_{HI} = 1, w_{VI} = 0 & \text{en-}route, \text{terminal}, NPA \\ w_{DOP} + w_{HI} + w_{VI} = 1 & \text{APV} \end{array}\right\} \tag{16.17}$$

It should be noted that the threshold values are utilized to normalization the effect of DOP and integrity loss. The PDOP loss and slope loss with removing the ith satellite is calculated as:

$$\Delta xDOP = \begin{cases} xDOP_i - xDOP_o & \Delta xDOP_i < Th_{\Delta xDOP} \\ \text{Inf} & \Delta xDOP_i \geq Th_{\Delta xDOP} \end{cases} \tag{16.18}$$

$$xSlope = \begin{cases} xSlope_i & xSlope_i < Th_{xSlope} \\ \text{Inf} & xSlope_i \geq Th_{xSlope} \end{cases} \tag{16.19}$$

The satellite i with minimum cost function J_i will be removed as causing least effect on the positioning accuracy and integrity. The selection/removing process will not stop until the number of satellite with specific cost function value is less than 5, i.e., the minimum number of satellites to maintain the ability of integrity. The flowchart of proposed weighted satellite selection algorithm is shown in Fig. 16.1. It is should be noted that the non-central parameter λ could be obtained with the off-line with the input of number of satellite, P_{fa} and P_{md}, otherwise will be negative for the time consumption performance and satisfying the time-to-alarm (TTA) requirement.

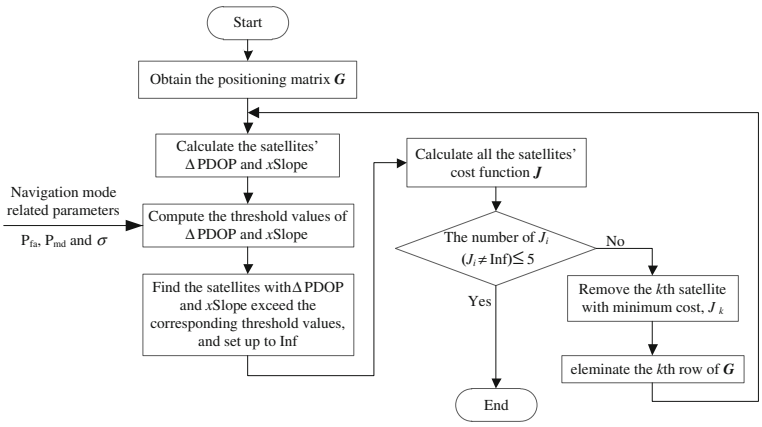


Fig. 16.1 Flowchart of weighted satellite selection algorithm

As seen from Fig. 16.1, the proposed algorithm is snapshot and could ultimately incorporate with the integrity monitoring and positioning process in real-time.

16.4 Evaluation and Discussion

The proposed weighted satellite selection algorithm (WSSA) was validated by comparing with other heuristic algorithms, i.e., MVA and QOA, in terms of integrity and accuracy. These heuristic algorithms were chosen with the purpose of demonstrating the edge of the proposed algorithm and due to their popularity in satellite selection research.

16.4.1 Integrity Performance Based on Simulation

In order to have a comprehensive integrity evaluation of the WSSA, the global availability simulation was carried out with the simulated GPS and GLONASS constellation. The users were distributed evenly with the interval of 5°. The simulation time length is 12 h with the sample period of 30 s. The cut-off elevation angle is set up to 7.5°. As GPS and GLONASS-constellation based RAIM have difficulty in satisfying the availability of APV II, this simulation was implemented at NPA and APV I respectively in order to show the edge of proposed algorithm in enhancing the availability level.

16.4.1.1 NPA Mode

The P_{fa} and P_{md} parameters are set up according to Table 16.1. The standard deviation of measurement error is set up to 15 m. The availability with using all satellites in-view is shown in Fig. 16.2a. It can be seen from Fig. 16.2a that the availability with using all satellites in-view is able to satisfy the predefined requirement.

As availability is the strictest parameter to evaluate the integrity performance and to be more extensive reveal the integrity performance with using different satellite selection algorithm, the availability ratio value between the algorithms to be tested (i.e., MVA, QOA and WSSA) $Availability_{proposed}$ and the one with using all satellite in-view (ASI) $Availability_{ASI}$ is defined as $\xi = Availability_{proposed}/Availability_{ASI}$. The ratio values from using MVA, QOA and WSSA were shown in Fig. 16.2b, c and d, respectively. As seen from Fig. 16.2, the availability of MDA and QOA are worse than the one of ASI, which means unable to satisfying the required navigation performance. However, the availability of WSSA can be improved compared with ASI. To furthermore reveal the origination of availability improvement due to WSSA, the change of parameter λ with according to different

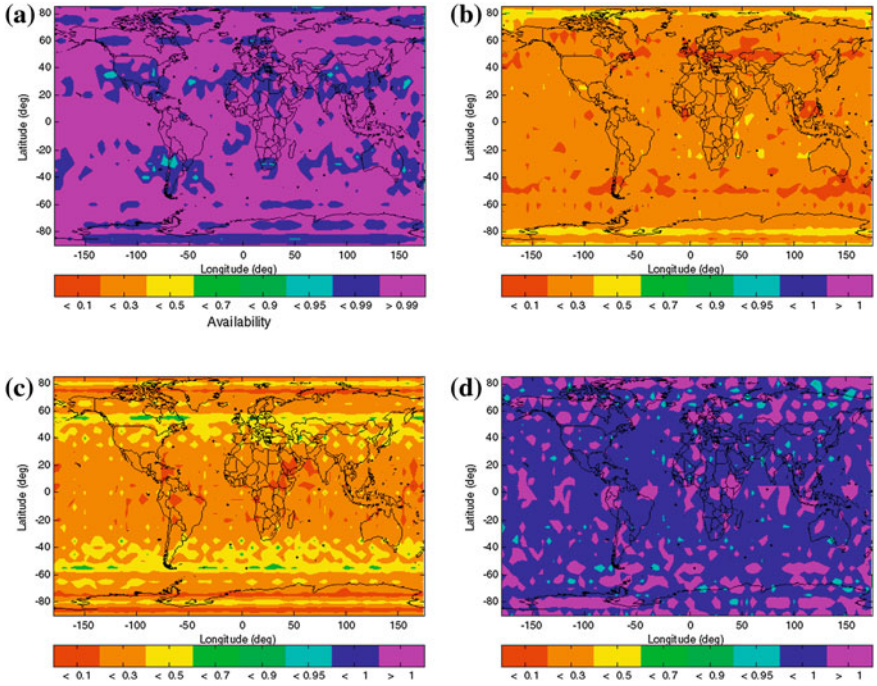


Fig. 16.2 Availability performance with using different satellite selection algorithms. **a** Availability with ASI. **b** ξ with MVA. **c** ξ with QOA. **d** ξ with WSSA

navigation phases is shown in Fig. 16.3. The reduction of satellites could lead the increasing *xSlope* in Eq. (16.4), which degrade the availability level, but also the decreasing of λ can be anticipated due to the reduction of satellites, as can be seen from Fig. 16.3. It is reasonable that the availability of WSSA can be anticipated since the *xSlope* increment can be compensated with the decrement of number of satellites.

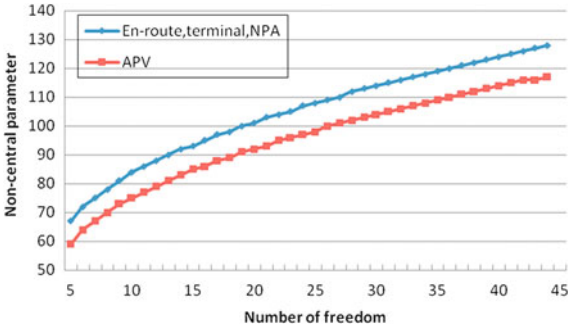


Fig. 16.3 The changing of λ under different navigation phases

16.4.1.2 APV I Mode

The availability performance with using different algorithms under APV I mode are shown in Fig. 16.4. The P_{fa} and P_{md} parameters are set up according to Table 16.1, and the stand deviation of measurement error is set up to 6 m. As shown in Fig. 16.4, it is difficult for ASI to satisfy the availability requirement of APV I. The availability of MVA and QOA are much worse than the ASI since MVA and QOA pay no attention to integrity and the *xPL* increase with the reduction of satellites. The availability of WSSA is also unable to satisfy the RNP of APV I, which mainly due to the increment of *xSlope* cannot be compensated by the decrement of λ.

16.4.2 Real-Time Performance Evaluation

In this section, the proposed algorithm is demonstrated with using real world GPS and GLONASS data logged on 11 November 2012 in Beijing, China. The total data collecting time-length is 5,139 s with the sample frequency of 1 Hz. The accuracy reference of the collected data was obtained from the differential carrier-phase positioning based RTK.

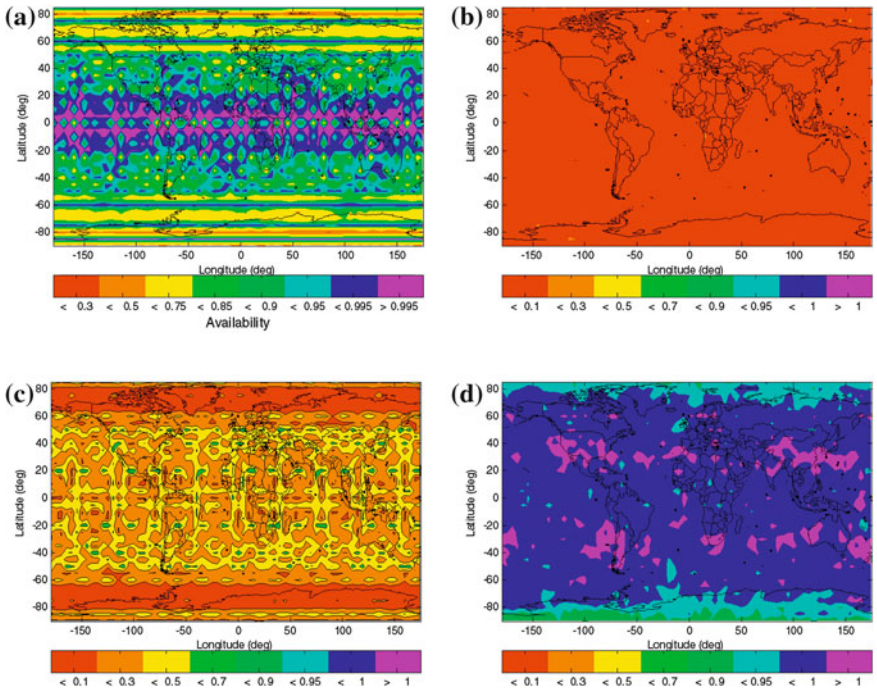


Fig. 16.4 Availability performance under APV I. **a** Availability with ASI. **b** ξ with MVA. **c** ξ with QOA. **d** ξ with WSSA

16.4.2.1 Accuracy and Integrity Evaluation

Since it has been shown in previous section that MVA and QOA have limitation in satisfy the availability requirement of NPA mode, the number of satellites and GDOP value after satellite selection by WSSA, MVA and QOA in the NPA mode are therefore shown in Fig. 16.5. The MVA is set up to select five satellites to maintain the integrity ability. As seen from Fig. 16.5, the number of satellites from WSSA and QOA is the same, but the satellite geometry of the two algorithms can be different, which can be proved by PDOP values of the two algorithms. The PDOP of WSSA is larger than that of QOA as the integrity consideration. Although MVA aims at minimizing PDOP loss, but the minimum is based on five satellites, which is less than those of WSSA and QOA, it is therefore the PDOP value of MVA is larger than QOA and WSSA.

In order to evaluate the accuracy performance of WSSA, the 3DRMS positioning error of three algorithms is shown in Fig. 16.6. It is shown that there is one jump exist in the QOA based positioning solution, which is an obviously position failure due to the loss of critical satellites [15]. With using the proposed WSSA, the positioning failure can be avoided. As a result, it can be seen that the effect of satellite selection with the consideration of integrity on improving accuracy is little but positive on avoiding the positioning failure.

To demonstrate the edge of WSSA in satisfying the integrity requirement, the HPL computed by WSSA, ASI, MVA and QOA are shown in Fig. 16.7. For the most part of experiment, the WSSA based HPL is less than the one of MVA and QOA, which means a better availability can be obtained. It is noted that WSSA based HPL is worse than the one without satellite selection, but in some cases a smaller WSSA based HPL can also be achieved due to the benefit from the reduction of number of satellites.

The 3DRMS95 % positioning error and availability with using different satellite selection algorithms is shown in Table 16.2 in order to make a numerical comparison and express the superior performance of WSSA. It can be seen that the positioning accuracy is not the best of the three algorithms, there is a remarkable superior performance in terms of availability, i.e., the availability of WSSA is

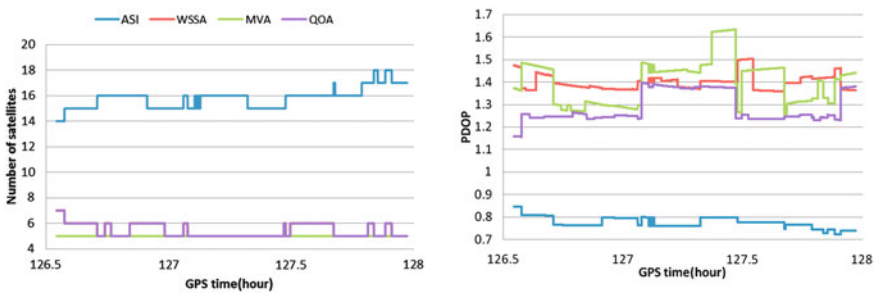


Fig. 16.5 The number of satellites (*Left*) and the PDOPs (*Right*) computed by three testing algorithms

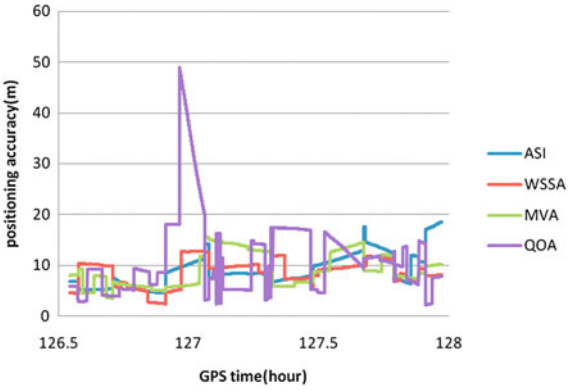


Fig. 16.6 The 3DRMS positioning error

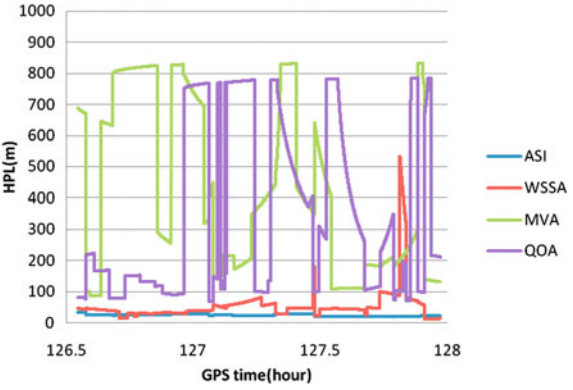


Fig. 16.7 The HPLs

much higher (47–59 %) relative to the QOA and MVA. It is because the proposed WSSA take account of integrity loss due to the reduction of satellites.

16.4.2.2 Time Performance

This section presents the computation time of each testing algorithm along the real-data test process. The CPU of the testing computer is 3.29 GHz with Intel® Core i3. Figure 16.8 shows the time consumption performance of MVA, QOA and WSSA. It is noted that QOA and WSSA select the same number of satellites in order to make convenient performance comparison.

As seen from Fig. 16.8, QOA have the less time consumption than the MVA and WSSA in the most part of testing process, while the time consumption of MVA linear increasing with the number of all in-view satellites n due to the large possible combinations (C_n^5). Although WSSA higher computation load than QOA, and the one of MVA in some cases, but the mean value of computation time (0.091 s) is medium, compared with the QOA (0.070 s) and MVA (0.124 s). It is

Table 16.2 Statistics of positioning error and availability of different algorithms

	ASI	MVA	QOA	WSSA
3DRMS95 %(m)	10.4	11.8	36.7	12.9
Availability (%)	100	39.27	51.43	99.02

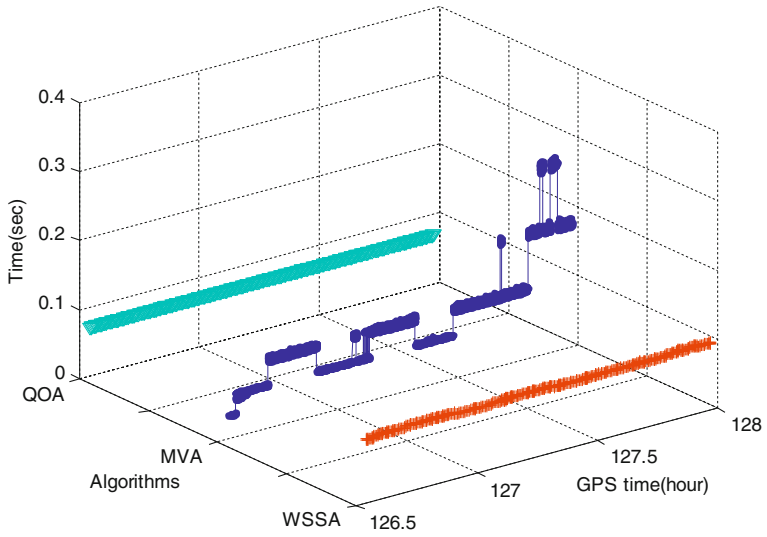


Fig. 16.8 Time performance of three testing algorithms

noted that the absolute values of time consumption of the three algorithms do not stand for the actual time consumption in one general commercial receiver, the time consumption performance indicates that the proposed WSSA is acceptable for real-time applications.

16.5 Conclusion

In this paper, a weighted satellite selection algorithm taking account of DOP and integrity is developed, implemented and tested. The main purpose of WSSA is to improve the overall performance of integrated GNSS receiver by maintaining acceptable positioning accuracy and minimal integrity loss. The threshold number of satellites be selected in WSSA is adaptively adjusted according to the RNP of the navigation phases without creating overload computation burden.

The evaluation results from real-world experiment and global integrated GPS/GLONASS simulation indicated that WSSA provides the optimal performance in positioning accuracy and availability level while requiring acceptable computation time, which will not be the major obstacle to real-time performance.

It can be concluded that the proposed algorithm can be option in improving the satellite selection performance. It should be noted that the proposed WSSA have the ability in detecting and excluding the fault satellite signal, the further work should be tested in the real/artificial interfere cases.

Acknowledgments The authors gratefully acknowledge the joint support from National High Technology Research and Development Program 863 (Grant No. 2012AA120505) and the Open Fund of Key Laboratory of Precision Navigation and Technology, National Time Service Center, CAS (2012PNTT15). This paper is provided for technical discussion, not represents any official standpoints.

References

1. Wu XL, Zhou JH, Wang G et al (2012) Multipath error detection and correction for GEO/IGSO satellites. Sci China-Phys Mech Astron 55(7):1297–1306
2. Roogpiboonsopit D, Karimi H (2009) A multi-constellation satellite selection algorithm for integrated global navigation satellite systems. J Intell Transp Syst: Technol, Plann Oper 13(3):127–141
3. Park CW (2001) Precise relative navigation using augmented CDGPS. PhD dissertation, Stanford University
4. Blanco-Delgado N, Nunes F (2010) Satellite selection method for multi-constellation GNSS using convex geometry. IEEE Trans Veh Technol 59(9):4289–4297
5. Parkinson B, Spilker J (1996) Global positioning system theory and applications (volumes II)
6. Wang J, Ober P (2009) On the availability of fault detection and exclusion in GNSS receiver autonomous integrity monitoring. J Navig 62(2):251–261
7. Enge P (1999) Local area augmentation of GPS for the precision approach of aircraft. Proc IEEE 87(1):111–132
8. Ochieng W, Sauer K, Walsh D, Brodin G, Griffin S, Denny M (2003) GPS integrity and potential impact on aviation safety. J Navig 56(1):51–65
9. Brown A, Sturza M (1990) The effect of geometry on integrity monitoring performance. In: Proceedings of the 46th annual meeting of the institute of navigation, Atlantic City, NJ, pp 121–129
10. Jin L, Huang Z, Li R, Ma Y (2009) Study on fast satellite selection algorithm for integrated navigation. Acta Electronica Sinica 37(9):1931–1936
11. RTCA (2004) Minimum aviation system performance standards for local area augmentation system (LAAS). Prepared by SC-159
12. Ren K, Yang L, Feng Y (2009) Comparison of two RAIM availability analysis approaches. J Telemetry, Tracking Command 30(2):13–16
13. Brown RG (1992) A baseline GPS RAIM scheme and a note on the equivalence of three RAIM methods. Navigation 39(3):301–316
14. Ciollaro M (2007) GNSS multisystem integrity for precision approaches in civil aviation. PhD dissertation
15. Li L, Zhao L, Ding J, Gao S (2011) A new inflation integrity monitoring algorithm for improving availability of LAAS signal-in-space. Acta Aeronautica et Astronautica Sinica 32(4):664–671

Chapter 17
Integrity Analysis of GPS Signal-in-Space Error

Dongjin Wang, Zhigang Huang and Rui Li

Abstract The space segment failure is one of the main sources of system integrity risk. This paper proposes an integrity analysis method to analyze the integrity risk of signal in space. Firstly, the original data should be preprocessed to calculate measured signal-in-space error. Secondly, the regularity of the errors is estimated by the spectrogram, followed by fitting orbit error with the least squares method based on Fourier series and AR model and fitting clock error with AR model. In order to detect the abnormal values of the signal-in-space errors roughly, the paper set the condition that when the difference between the original error and the fitted values is more than 4.42 times the standard deviation of fitted values, the abnormal data may exist. Then it's necessary to identify specific abnormal factors which conclude satellite unhealthy, duplicate broadcast ephemeris and ephemeris missing to statistics the according abnormal rate. The simulation results of the measured data of different type in-orbit satellite which is from 2002 to 2011 show that the abnormal rate of signal in space is 10^{-4}/h, which meet the index that three satellites may fault every year, therefore the analysis method is verified to be effective.

Keywords Signal-in-space error · Integrity risk · Fitting model · Abnormal rate

17.1 Introduction

The space segment is one of the main sources of system integrity risk, mainly covers a series of anomalies caused by satellite itself. Satellite anomalies will lead to wrong navigation signal or the fact navigation information can't correctly reflect

D. Wang (✉) · Z. Huang · R. Li
School of Electronic and Information Engineering of Beihang University, Beijing, China
e-mail: wangdj1991@bupt.edu.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_17,

the real situation, which may cause mutational pseudo range. Therefore, establishing effective integrity analysis method and determining the characteristics and causes of the space segment's abnormalities are the premise of system integrity monitoring process and have great significance.

From the view of abnormal rate of measured signal-in-space errors, the paper proposes a signal-in-space integrity analysis method. On the basis of the regularity of measured signal-in-space errors of navigation system, this article establish fitting model to detect the potential abnormal values and indentify the characteristics and causes of abnormal factors and statistics the according abnormal rate. At last, the method is verified to be effective based on nearly 10 years' measured data of GPS.

17.2 Signal-in-Space Error

17.2.1 Original Data Extraction and Processing

Signal-in-space error was defined as any errors related to satellite segment, mainly satellite orbit errors and clock errors. The traditional calculating method is reading orbit and perturbation term parameters from RENIX format broadcast ephemeris and calculating orbit and clock information in Earth Centered, Earth Fixed(ECEF) coordinate system based on the formula of interface document, and then comparing with that of precise ephemeris at the same epoch.

In this paper, the type of broadcast ephemeris is brdcdd.yy, and the type of precise ones providing by International Geodynamic Service (IGS) is SP3-type. Precise ephemeris data conclude satellite orbit and clock data in International Terrestrial Reference Frame every 15 min synchronized to GPS time without considering the coordinate system conversion error [1].

In order to effectively detect the abnormal values, the epoch interval is set to 1 min so that it's necessary to interpolate precise ephemeris. This paper use the 12-order Lagrangian interpolation and locate the interpolation nodes in the center at the same time, that means the six groups before and after the nodes are seen as an interpolation interval.

17.2.2 The Variation Regulation of Signal-in-Space Errors

The GPS has 32 in-orbit satellites which can be divided into IIA, IIR, IIR-M and IIF according to satellite clock. Tacking the fact that different types of satellites have different performance and launch time so that may affect the accuracy of the according signal-in-space errors into account, this article will analyze these four types of satellites. Due to space limitation, Fig. 17.1 shows only the simulation results of the signal-in-space of IIF-1 satellite and those of other existing types of satellites from March 23, 2011 to March 29, 2011.

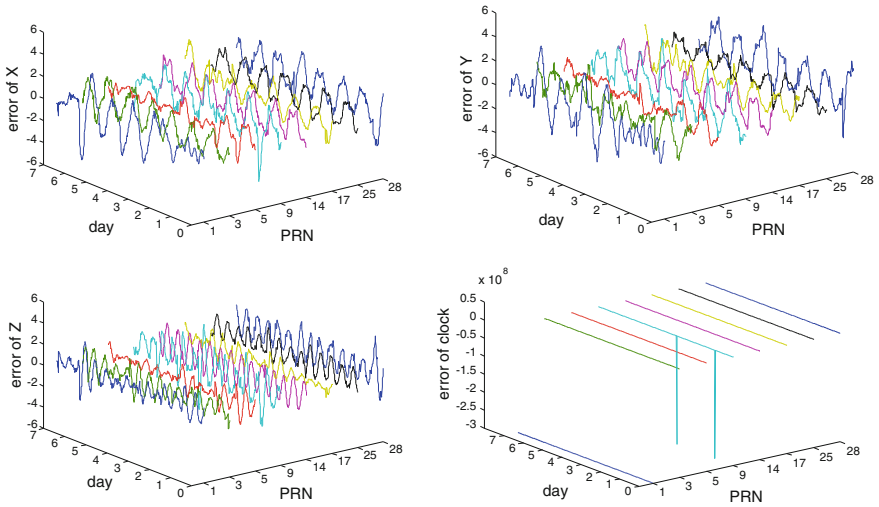


Fig. 17.1 Signal-in-space errors of different types of satellites from March 23, 2011 to March 29, 2011

The orbit errors are within $[-4, 4]$ meters varying periodically and have nothing to do with the satellites' type; while the clock errors are not periodic. We can also see that the clock errors of satellite PRN 1 and PRN 9 are abnormal.

17.3 The Signal in Space Integrity Analysis Method

17.3.1 Crude Detection of Abnormal Values

Considering the result in the previous section that signal-in-space errors are regular, detecting abnormal values by establishing an effective fitting model can be achieved.

17.3.1.1 Fitting Model of Orbit Errors and Clock Errors

Unsmooth cyclical sequence may cause spurious regression, so that model can be established by the theoretical and analytical tools of the stationary process based on the sequence that already remove periodic terms. Then the spectrum of orbit errors is given to determine the cycle. Due to space constraints, Fig. 17.2 shows the spectrum of signal-in-space errors of satellite PRN 28 from March 27, 2011 to April 2, 2011.

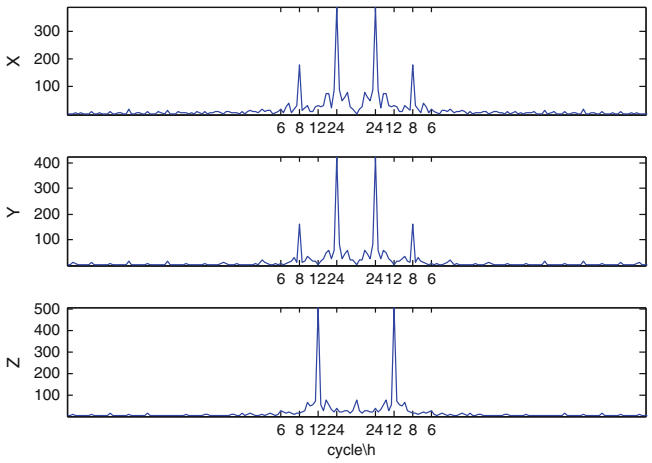


Fig. 17.2 Spectrum of orbit errors of satellite PRN 28 from March 27, 2011 to April 2, 2011

It's obvious that the main periodic terms of X and Y direction of orbit errors are 24 and 8 h; while that of Z direction is 12 h. We can use the following equation to express the orbit errors.

$$\Delta R = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} 1 & \sin(\frac{t}{T_1}) & \cos(\frac{t}{T_1}) & \sin(\frac{t}{T_2}) & \cos(\frac{t}{T_2}) \\ 1 & \sin(\frac{t}{T_1}) & \cos(\frac{t}{T_1}) & \sin(\frac{t}{T_2}) & \cos(\frac{t}{T_2}) \\ 1 & \sin(\frac{t}{T_3}) & \cos(\frac{t}{T_3}) & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} + \begin{bmatrix} \Delta\Delta x \\ \Delta\Delta y \\ \Delta\Delta z \end{bmatrix} = H \cdot A + \Delta\Delta R \tag{1.1}$$

where t is epoch, A_i is the according cycle coefficient, T_1, T_2 and T_3 are 24, 8 and 12 h, respectively, $[\Delta\Delta x, \Delta\Delta y, \Delta\Delta z]^T$ is the non-periodic term residuals of X, Y and Z direction of orbit errors, respectively. The parameter vector is estimated by the method of least squares as shown in the following equation.

$$\hat{A} = \left(H^T H\right)^{-1} H^T \cdot \Delta R \tag{1.2}$$

The above non-periodic residual sequence and the clock errors can be approximately regarded as stationary random sequence, thus AR fitting model is used. Taking the X direction of orbit errors for example, the following equation can [2] be get.

$$\Delta\Delta x(t_i) = \sum_{j=1}^{n} B_j \cdot \Delta\Delta x\left(t_{i-j}\right) + \Delta\Delta\Delta x_i \tag{1.3}$$

where B_j is the according coefficient, n is order of AR model, and $\Delta\Delta x_i$ is the residuals. The matrix forms of expression and the parameter estimated by the least squares method are as shown in the following equations.

$$X = \begin{bmatrix} \Delta\Delta x(t_{i+1}) \\ \Delta\Delta x(t_{i+2}) \\ \vdots \\ \Delta\Delta x(t_{i+n}) \end{bmatrix} = \begin{bmatrix} \Delta\Delta x(t_i) & \Delta\Delta x(t_{i-1}) & \cdots & \Delta\Delta x(t_1) \\ \Delta\Delta x(t_{i+1}) & \Delta\Delta x(t_i) & \cdots & \Delta\Delta x(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Delta\Delta x(t_{n-1}) & \Delta\Delta x(t_{n-2}) & \cdots & \Delta\Delta x(t_{n-i}) \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_i \end{bmatrix} + \Delta X$$
$$= D \cdot B + \Delta X \tag{1.4}$$

$$\hat{B} = \left(D^T D\right)^{-1} D^T \cdot X \tag{1.5}$$

When AR model is used, it must take the order into account and the criterion commonly used is final prediction error (FPE), just as the following equation.

$$\text{FPE} : \frac{\text{N} + \text{n}}{\text{N} - \text{n}} * \text{var} \tag{1.6}$$

Figure 17.3 shows the correspondence between FPE criterion and order. When the order is set to 2, the fitting effect of signal-in-space errors of satellite PRN 28 from March 27, 2011 to April 2, 2011 is shown as Fig. 17.4.

In order to quantitatively measure the accuracy of the fitting model, the concept of the improvement rate of mean square deviation is introduced. Taking X direction of orbit errors for example, we set the variance before and after fitting is σ_1 and σ_2, respectively, thus the improvement rate can be represented as follow (Table 17.1).

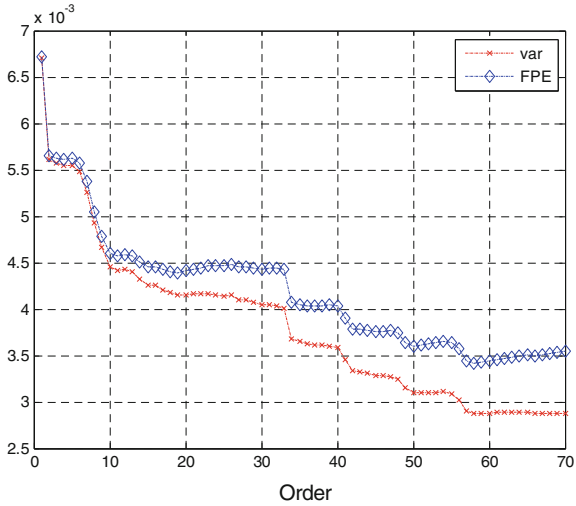


Fig. 17.3 Correspondence between FPE and order

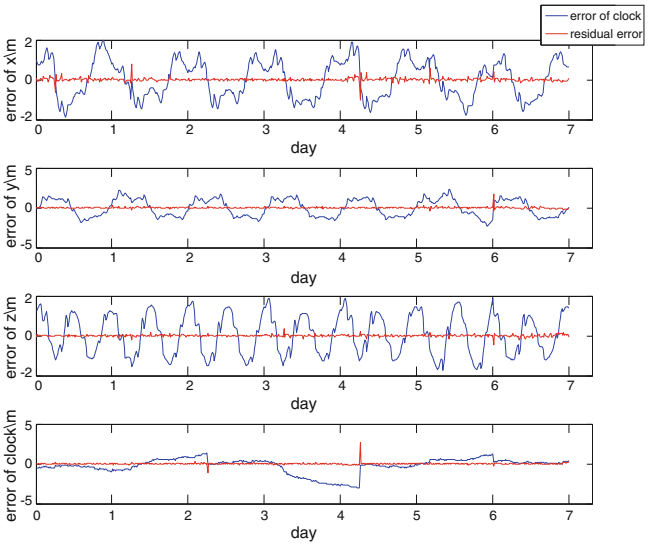


Fig. 17.4 Fitting effect of signal-in-space error

Table 17.1 Correspondence between variance average improvement and order

Order	R_x (%)	R_y (%)	R_z (%)	R_clock (%)
2	89.75	91.66	92.09	85.95
8	90.07	92.18	92.99	86.13
25	90.24	92.51	93.74	86.60
50	90.38	93.44	94.32	87.16
60	90.42	93.65	94.42	88.32

$$\Delta\sigma = \frac{\sigma_1 - \sigma_2}{\sigma_1} \tag{1.7}$$

Table 17.1 shows the correspondence between variance average improvement and the order, so that we can draw a conclusion, that the corrective effects of 2-order AR model to orbit errors and clock errors are about 90 and 86 %. Considering the middleweight of orbit and clock errors are 2 and 4 m, thus the fitting residuals are about 0.2 and 0.5 m, respectively. The satellites' clock transition leads to reduction of the improvement rate of mean square deviation, however, clock transition can't be predicted in practice as well, so we can draw a conclusion that the above fitting model can meet basic needs. With increasing order of AR model, the correct effect increase gradually, but the amount of computation will significantly improve. Thus, with comprehensive consideration of computational cost and correction effect, the order of AR model is set to 2.

17.3.1.2 Abnormal Values Threshold Set

Signal-in-space range error (SISRE) is defined to indicate the accuracy of signal-in-space errors, just as the following equation.

$$SISRE = \sqrt{(R - clock)^2 + \frac{1}{49}(A^2 + C^2)} \tag{1.8}$$

where R, A and C is the radial, along-track and cross-track values that orbit errors vector projected onto the satellite inertial coordinate system, respectively.

Known from the literature [3], the satellite SISRE reduced to 2 m by 5 m from 2002 to 2011. Therefore, this article sets the SISRE corresponding to the fitting values of signal-in-space errors as the mean, and defines when the difference between actual SISRE and the fitting one exceeds 4.42 times the standard deviation of the fitting SISRE, the data on the epoch is considered to be abnormal.

17.3.2 Abnormality Rate of Signal in Space

It's necessary to indentify abnormal factors and their causes, and figure out the corresponding abnormal rate, in order to determine whether the abnormalities detected in the previous section are caused by the exception of signal-in-space itself.

Firstly, the paper simulates the signal-in-space errors of satellite PRN 3 and PRN 28 from January 1, 2002 to September 30, 2011, epoch interval of 1 min. Due to space limitation, Fig. 17.5 only shows satellite PRN 28's signal-in-space errors in September 26, 2009. In addition to clock errors, orbit errors can also be abnormal.

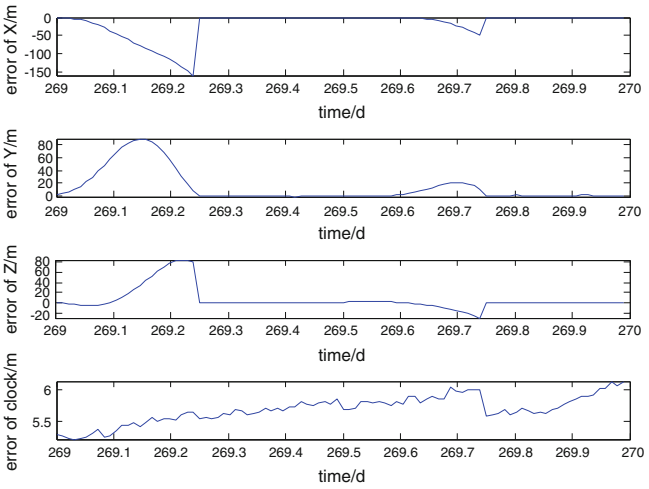


Fig. 17.5 Sigal-in-space error of satellite PRN28 on September 26, 2009

There are four potential factors will lead to abnormal satellite signal-in-space errors, respectively, such as broadcast ephemeris missing, repeat ephemeris, satellite unhealthy and abnormal clock errors correction of precise ephemeris.

1. Broadcast ephemeris provide a set of orbital parameters every 2 h, however the actual ephemeris may miss. Taking satellite PRN 3 for example, the broadcast ephemeris on January 1, 2011 contains only two sets of parameters, as shown in Fig. 17.6.

Orbit and clock errors at any epoch are extrapolated by the formulas provided by the interface documentation based on the orbital parameters of broadcast ephemeris. Therefore it's inevitably to appear coordinate values jump at alternating moment, values' jump due to extrapolated time of 2 h are just around a few centimeters so that it can be negligible. But when the broadcast ephemeris is missing, the accuracy of orbit and clock errors will sharply decline with the extension of extrapolation time, and even lead to hundreds of meters abnormal error.

2. The broadcast ephemeris exist the condition that the navigation message conclude the whole point and non-whole point of orbit parameters. Figure 17.7 shows the ephemeris of satellite PRN 28 at 9:59:44 and 10:0:0 on October 11, 2010.

The orbit and clock information from 10 o'clock to 12 o'clock are extrapolated by the orbital parameters at 10 o'clock, however, the parameters may be wrong, opposite the parameters at 9:59:44 can be trusted. Therefore, signal-in-space errors may jump and be abnormal.

3. The broadcast ephemeris provide parameter 'svhealth' to mark satellite health status so that users can avoid using wrong orbital information because of satellite attitude adjustment or test.
4. The actual precise ephemeris files may have the condition that satellite clock corrections are wrong obviously, such as the values of 999999.999999 which

```
     2             NAVIGATION DATA                         RINEX VERSION / TYPE
CCRINEXN V1.6.0 UX  CDDIS               10-AUG-11 21:42     PGM / RUN BY / DATE
IGS BROADCAST EPHEMERIS FILE                                COMMENT
    0.7451D-08 -0.1490D-07 -0.5960D-07  0.1192D-06          ION ALPHA
    0.9216D+05 -0.1147D+06 -0.1311D+06  0.7209D+06          ION BETA
   -0.372529029846D-08-0.133226762955D-13    61440     1617 DELTA-UTC: A0,A1,T,W
    15                                                      LEAP SECONDS
                                                            END OF HEADER
 3 11  1  1  0  0  0.0 0.656384043395D-03 0.522959453519D-11 0.000000000000D+00
    0.650000000000D+02 0.515625000000D+01 0.546487033404D-08 0.186044562888D+01
    0.726431608200D-07 0.138873044634D-01 0.706873834133D-05 0.515362539673D+04
    0.518400000000D+06-0.175088644028D-06 0.590814834267D+00-0.279396772385D-06
    0.928201164790D+00 0.227500000000D+03 0.106799018775D+01-0.836570546170D-08
    0.444304232294D-09 0.100000000000D+01 0.161600000000D+04 0.000000000000D+00
    0.200000000000D+01 0.000000000000D+00-0.465661287308D-08 0.650000000000D+02
    0.511218000000D+06 0.400000000000D+01 0.000000000000D+00 0.000000000000D+00
 3 11  1  1  2  0  0.0 0.656421761960D-03 0.522959453519D-11 0.000000000000D+00
    0.740000000000D+02 0.868750000000D+01 0.536165201126D-08 0.291064183201D+01
    0.394880771637D-06 0.138868535869D-01 0.675395131111D-05 0.515362636757D+04
    0.525600000000D+06 0.707805156708D-07 0.590756054218D+00-0.145286321640D-06
    0.928205297534D+00 0.229593750000D+03 0.106801035115D+01-0.811962408420D-08
    0.410731993341D-09 0.100000000000D+01 0.161600000000D+04 0.000000000000D+00
    0.200000000000D+01 0.000000000000D+00-0.465661287308D-08 0.740000000000D+02
    0.518418000000D+06 0.400000000000D+01 0.000000000000D+00 0.000000000000D+00
```

Fig. 17.6 Broadcast ephemerides missing

```
28 10 10 11  9 59 44.0-0.125728547573D-06 0.159161572810D-11 0.000000000000D+00
    0.500000000000D+01-0.124062500000D+02 0.440268338944D-08-0.184937198417D+00
   -0.584870576858D-06 0.162667608820D-01 0.972114503384D-05 0.515368316078D+04
    0.122384000000D+06-0.143423676491D-06 0.108806372058D+01 0.160187482834D-06
    0.971571153084D+00 0.199562500000D+03-0.194477696498D+01-0.817855495536D-08
    0.421446126351D-09 0.100000000000D+01 0.160500000000D+04 0.000000000000D+00
    0.200000000000D+01 0.000000000000D+00-0.107102096081D-07 0.500000000000D+01
    0.119220000000D+06 0.000000000000D+00 0.000000000000D+00 0.000000000000D+00
28 10 10 11 10  0  0.0-0.128988176584D-06 0.159161572810D-11 0.000000000000D+00
    0.850000000000D+02-0.124062500000D+02 0.440268338944D-08-0.182603610013D+00
```

Fig. 17.7 Duplicate broadcast ephemerides

```
P  3   8275.490417 -23133.547294  -9927.550085      8.081872
P  3   9110.980912 -23754.894020  -7365.197653      8.085043
P  3   9759.678349 -24163.226939  -4675.039308      8.087032
P  3  10242.796788 -24329.518410  -1903.657551 999999.999999
P  3  10587.120478 -24231.125371    900.828647 999999.999999
P  3  10823.766553 -23852.761308   3689.622245      8.098808
P  3  10986.776418 -23187.192360   6414.125967      8.101725
```

Fig. 17.8 Abnormal clock error corrections of precise ephemeris

represent the clock corrections are unknown. Figure 17.8 shows the precise ephemeris of satellite PRN 3 on May 8, 2003.

Precise ephemeris only as a true value reference in signal-in-space integrity analysis, we don't consider the exception caused by precise ephemeris' abnormal clock corrections. Broadcast ephemeris is forecast ephemeris estimated by master station method based on five stations' observational data by Kalman filtering method. Different monitor stations will record the navigation message of satellite in view, therefore duplicate broadcast ephemeris and ephemeris missing may be caused by the process of handling monitoring data, rather than signal-in-space itself. Considering that we can only obtain after-processing data, the deduced abnormal rate must be greater than the actual abnormal rate of signal in space. The above abnormal rate is defined as duration time of abnormity divided by the total statistical time in this article.

17.4 Integrity Analysis Conclusion of Measured Signal in Space

The signal-in-space integrity of different types of satellites from 2002 to 2011 is simulated in order to verify the effectiveness of the integrity analysis method. The abnormal rates of signal-in-space each year are shown in Fig. 17.9. Due to space limitations, only the abnormal rate of IIA-type satellite PRN 3, IIR-type satellite PRN 28, IIR-M-type satellite PRN 29 and IIF-type satellite PRN 25 are shown. The simulation start time of different satellites is different because of different launch time and valid time as shown in Table 17.2. Figure 17.10 only shows the annual percentage of abnormal factors of satellite PRN 3 due to space limitation (Fig. 17.9, Table 17.2).

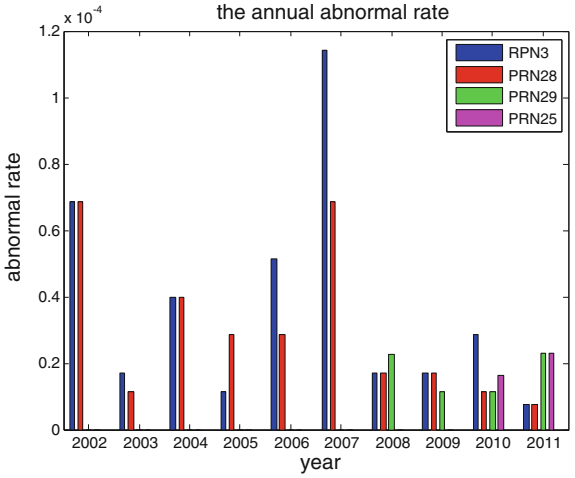


Fig. 17.9 The abnormal rate of signal in space each year

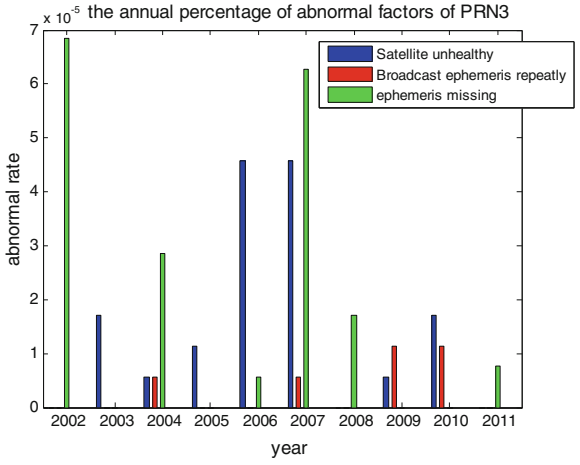


Fig. 17.10 The annual percentage of abnormal factors of PRN3

It's easily noted that the abnormal rate of the above four types of satellites is about 10^{-4}/h. In view that broadcast ephemeris missing and repeated ephemeris may be caused by the process of monitor station data processing, the actual abnormal rate of signal in space itself must be lower than 10^{-4}/h. The above conclusion meets the indicators of GPS that there must be up to three failure satellites annually, therefore the integrity analysis method is effective.

We can also draw some other conclusions. The abnormal rate of signal-in-space errors reduce with the growth of useful life gradually; main abnormal factors are satellite unhealthy and ephemeris missing; the abnormal rate of IIF-type satellites is lower than the other three types of satellites, and the clock performance improves significantly.

Table 17.2 Valid simulation time

	PRN 3	PRN 28	PRN 29	PRN 25
Start time	2002.1.1	2002.1.1	2008.1.1	2010.8.25
End time	2011.9.30	2011.9.30	2011.9.30	2011.9.30

17.5 Conclusion

This paper proposes an integrity analysis method to analyze the integrity risk of signal in space. Firstly, the regularity of signal-in-space errors after pretreatment is analyzed by the spectrogram, followed by fitting orbit error with the least squares method based on Fourier series and AR model and fitting clock error with AR model, in order to detect the abnormal values of the signal-in-space errors roughly, and identifying specific abnormal factors to statistics the according abnormal rate. At last, taking the measure signal in space of different types of satellites from 2002 to 2011 for example, the paper simulates the abnormal rate. The simulation results show that the abnormal rate of the above four types of satellites is about 10^{-4}/h. Excluding the exception caused by monitor stations, the actual abnormal rate of signal in space itself must be lower than 10^{-4}/h. The above conclusion meets the indicators of GPS that there must be up to three failure satellites annually, therefore the integrity analysis method is effective. Some other conclusions can also be obtained. The abnormal rate of signal-in-space errors reduces with the growth of useful life gradually; main abnormal factors are satellite unhealthy and ephemeris missing; the clock performance of IIF-type satellites is more excellent than the other three types. It can be expected that the future navigation systems enhance the data processing capabilities of monitor stations and launch new type satellites with more excellent performance to future reduce the integrity risk.

Acknowledgments The authors acknowledge the National Basic Research Program of China (Grant No. 2010CB731805) grants.

References

1. Cui X (2006) Comparisons of two kinds of GPS broadcast ephemeris parameter algorithms. Chin J Space Sci 26(5):282–387
2. Wende H, Wei W, Xiaoning X (2010) Frequency domain analysis and forecasting model of navigation satellite broadcast ephemeris error. China Space Sci Technol 6(3):12–18
3. Liang H, Gao GX (2011) Todd walter. Statistical characterization of GPS signal-in-space errors. In: Proceedings of ION GPS

Chapter 18
Performance Analysis of COMPASS/GPS Code Differential Positioning

Junyi Xu, Yuanxi Yang, Jinlong Li, Haibo He and Hairong Guo

Abstract The local navigation system, BeiDou 2 or COMPASS has been successfully established lately. By using real data of several baselines, the accuracy of COMPASS/GPS differential positioning under single and dual frequency situation was analysed to evaluate the expecting performance of COMPASS/GPS ground based augmentation system (GBAS). Besides, three different smoothing methods are used. The results show that, the differential position accuracy of COMPASS is lower than that of GPS, and it can be significantly improved by a long time smoothing, the divergence free smoothing is suggested to be used when dual frequency signals is available. And the precision and availability of COMPASS/GPS based GBAS is superior to that of GPS single system.

Keywords COMPASS · GPS · GBAS · Carrier phase smoothing pseudorange

18.1 Introduction

The global navigation satellite system (GNSS) is undergoing a flourishing development now. GPS has successfully expand its constellation to a 27-slot one in 2011, and its modernization is advancing steadily now. The GLONASS system has also been in full operation by then end of 2011, and it will provide triple frequency

J. Xu (✉) · J. Li
Institute of Surveying and Mapping, Information Engeering University, Zhengzhou 450052, China
e-mail: xujunyi-025@163.com

Y. Yang
National Key Laboratory of Geo-information Engineering, Xi'an 710054, China

H. He · H. Guo
Beijing Satellite Navigation Center, Beijing 100094, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_18,

CDMA signal by 2014 [1–3]. The BeiDou local navigation system or, COMPASS, has already successfully established (including 5 GEO, 5 IGSO, and 4MEO). The number of visible satellite will increase by multiple constellations, and the satellite geometry will be improved, which will further improve the accuracy, availability and integrity of satellite navigation [4].

However, the position accuracy that can be achieved by using BeiDou, GPS, and GLONASS is about 10 m because of the errors such as satellite clock, ionosphere delay and troposphere delay. This can not satisfy the requirement of applications such as precision approach, weapon accuracy test and precision-guided munitions. The differential technology is an effective way to enhance the accuracy and integrity of satellite navigation. So, scholars all over the world are researching on the ground based augmentation system (GBAS) to satisfy the requirement of these applications. The data processing strategy of GBAS by using dual frequency signals are studied by Konno [5]. Chen and Gan made a research on the GPS based GBAS [6, 7]. For GPS single system, usually pseudolite should be established to improve the geometry, and enhance the availability of navigation system. While for multi constellation situation, this may no longer be necessary any more. Thus, Hideki Y ea al made a research on the differential position accuracy of the GPS/GLONASS based GBAS [8, 9, 10].

There are seldom researches on COMPASS and COMPASS/GPS based GBAS since the BeiDou had just established recently. An elementary analysis way made by Shi et al. [11] who mainly focus on the Precise relative positioning.

Therefore, the position accuracy of COMPASS/GPS differential positioning is analyzed by using the real data. The results show that, the differential position accuracy of COMPASS is lower than that of GPS, and it can be significantly improved by a long time smoothing filter; the precision and availability of COMPASS/GPS differential positioning is superior to that of GPS single system.

18.2 Differential Positioning

There are two ways for differential positioning, broadcasting differential corrections and the original observable. The differential correction way was adopted here. First, the ground station computes the geometric range to the satellite,, from the reference receiver, [5]:

$$r_g^i = \left\| X^i - X_g \right\| \tag{18.1}$$

where, refers to the satellite position acquired from the navigation message, is the precise position of base station receiver.

Then, the differential correction of satellites:

$$\begin{aligned} e^i &= r_g^i - \rho_g \\ &= -c\left(\delta t_g - \delta t^i\right) - I_g^i - T_g^i - \varepsilon_g^i \end{aligned} \tag{18.2}$$

where, δt_g is the receiver clock error, δt^i the satellite clock error, I_g^i is the ionosphere delay, T_g^i is the troposphere delay, ε_g^i is the receiver noise and the multipath error.

By using the received differential corrections, the observables of rover station can be expressed as

$$\begin{aligned}\tilde{\rho}_u^i &= \rho_u^i + e^i \\ &= r_u^i + c\left(\delta t_u - \delta t^i\right) + \left(I_u^i - I_g^i\right) + \left(T_u^i - T_g^i\right) + \varepsilon_u^i - \varepsilon_g^i\end{aligned} \tag{18.3}$$

18.3 Carrier Smoothing

The code and carrier phase measurements of COMPASS and GPS on B1 (L1) and B2 (L2) can be simplified as

$$\begin{aligned}\rho_1 &= \mathrm{R} + I_1 + \varepsilon_{\rho 1} \\ \phi_1 &= \mathrm{R} - I_1 + N_1 + \varepsilon_{\phi 1} \\ \rho_2 &= \mathrm{R} + \frac{f_1^2}{f_2^2} I_1 + \varepsilon_{\rho 2} \\ \phi_2 &= \mathrm{R} - \frac{f_1^2}{f_2^2} I_1 + N_2 + \varepsilon_{\phi 2}\end{aligned} \tag{18.4}$$

where, ρ is the code phase measurement, ϕ is the carrier phase measurement, R is the geometric range from the receiver to the satellite, I_1 is the ionosphere delay on B1(L1), N is integer ambiguity, ε_ρ is the receiver noise and the multipath error on code phase measurement, ε_ϕ is the receiver noise and the multipath error on carrier phase measurement.

The carrier smoothing algorithms are usually used in GBAS to improve the position accuracy. The core of carrier smoothing is suppress the multipath and receiver noise and multipath error by using low pass filter(see Fig. 18.1). First, the code and carrier inputs are differenced, the difference value χ are smoothed, the output from the low-pass filter, $\hat{\chi}$, is recombined with the carrier phase measurement. The Hatch filter is used in this paper [12]

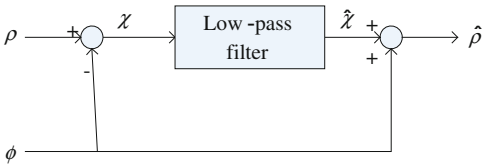


Fig. 18.1 Block diagram of carrier smoothing filter

18.3.1 Single Frequency Smoothing

For single frequency smoothing (SFCS), the difference of code and carrier measurement is

$$\chi_{SF} = \rho_1 - \phi_1 = \varepsilon_{\rho_1} + 2I_1 - N_1 \tag{18.5}$$

The smoothed difference and code measurement can be written as

$$\hat{\chi}_{SF} = \hat{\varepsilon}_{\rho_1} + 2\hat{I}_1 - N_1 \tag{18.6}$$

$$\hat{\rho}_{SF} = \hat{\chi}_{SF} + \phi_1 = R + \hat{\varepsilon}_{\rho_1} + \left(2\hat{I}_1 - I_1\right) \tag{18.7}$$

It can be see form Eq.(18.7) that, when the ionosphere delay varies or the smooth interval in inappropriate, the difference between the smoothed ionosphere delay $\hat{I}_1$ and the unsmoothed I_1 may be unacceptable, and this may yield another source of error, which will lead to the decadence of position accuracy.

18.3.2 Dual Frequency Smoothing

For dual frequency, normally the measurements are combined first, and then the code measurement is smoothed. At present, there are two smoothing methods, ionosphere free (IFree) smoothing and divergence free (DFree) smoothing [13]

When the IFree smoothing are used, the input measurement are:

$$\begin{aligned} \rho_{IF} &= \rho_{L1} - \frac{1}{\xi}(\rho_{L1} - \rho_{L2}) \\ \phi_{IF} &= \phi_{L1} - \frac{1}{\xi}(\phi_{L1} - \phi_{L2}) \end{aligned} \tag{18.8}$$

where, $\xi = 1 - \frac{f_1^2}{f_2^2}$.

The first order ionosphere effect is removed by IFree, however, the receiver noise and multipath error are amplified.

When the IFree smoothing are used, the input measurement are:

$$\begin{aligned} \rho_1 &= R + I_1 + \varepsilon_{\rho_1} \\ \phi_{DF} &= \phi_{L1} - \frac{2}{\xi}(\phi_{L1} - \phi_{L2}) \end{aligned} \tag{18.9}$$

The smoothed difference and code measurement can be written as

$$\chi_{DF} = \varepsilon_{\rho_1} - N_{LC,DF} \tag{18.10}$$

$$\hat{\rho}_{DF} = \hat{\chi}_{DF} + \phi_{DF} = R + \hat{\varepsilon}_{\rho_1} + I_1 \tag{18.11}$$

Compared with SFCS, the DFree smoothing will not yield new error; besides, the receiver noise and multipath error are not amplified.

18.4 Calculations and Analysis

The UB240-CORS dual frequency GPS/COMPASS receiver produced by Beijing Unicore Communications Incorporation was used for data collection. Three reference stations (record as A, B, C) which located in Beijing are selected, and the static test last 24 h, the sampling rate is 1 s. The baselines formed by the three stations are recorded as BC, AB, and AC, the baseline length is respectively 8.2, 35.4, and 42.2 km. The cut-off elevation is set as 10°, and the root mean square (RMS) error are used for position accuracy statistics.

Firstly, the position accuracy of unsmoothed single frequency differential positioning is calculated, the results can be seen in Figs. 18.2, 18.3 and Table 18.1. The sample rate is 10 s. Then, the carrier smoothing differential positioning for single and dual frequency is calculated. The data from 8:00 am to 20:00 pm are used, the original 1 s data were used, and the smoothing time is 100 s. The results can be seen in Tables 18.2, 18.3, and 18.4.

It can be seen from Table 18.1 that, without carrier smoothing,the position accuracy of COMPASS single frequency differential is about 2–4 m. The RMS of three baseline in Up component is 3.114, 2.167, and 2.642 m respectively, for GPS, the RMS in Up component is 1.116, 0.849, and 1.137 m; and the RMS reduced to 0.953, 0.705, and 0.918 m for COMPASS/GPS dual system. Figure 18.2 shows the position error of baseline AB in Up component. Figure 18.3 shows the three dimension position accuracy. It can be seen that performance of COMPASS is not as good as GPS, and the position accuracy of COMPASS/GPS is better than that of GPS.

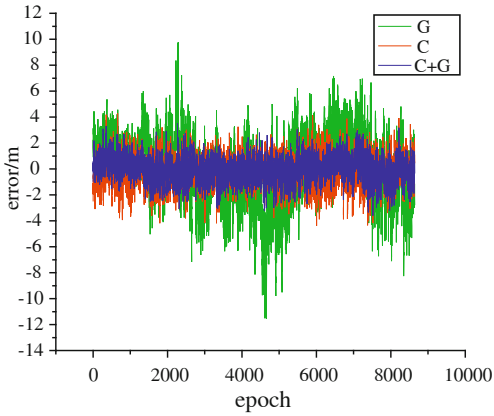


Fig. 18.2 *Vertical* error of baseline AB

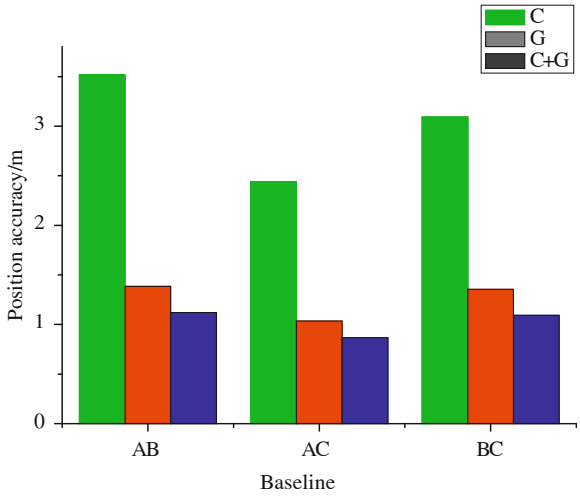


Fig. 18.3 D position accuracy comparison

Table 18.1 Accuracy comparison of unsmoothed single frequency differential position (m)

	AB			AC			BC		
	E	N	U	E	N	U	E	N	U
COMPASS	0.614	1.519	3.114	0.467	1.019	2.167	0.616	1.484	2.642
GPS	0.448	0.598	1.166	0.346	0.462	0.849	0.448	0.582	1.137
C + G	0.324	0.494	0.953	0.263	0.359	0.705	0.354	0.476	0.918

Table 18.2 Accuracy comparison of single frequency smoothed differential position (m)

	AB			AC			BC		
	E	N	U	E	N	U	E	N	U
COMPASS	0.555	1.257	2.884	0.463	1.045	2.018	0.543	1.180	2.223
GPS	0.210	0.315	0.677	0.207	0.308	0.566	0.224	0.302	0.618
C + G	0.214	0.301	0.652	0.194	0.255	0.563	0.239	0.267	0.586

Table 18.3 Accuracy comparison of ionosphere free smoothed differential position (m)

	AB			AC			BC		
	E	N	U	E	N	U	E	N	U
COMPASS	2.005	3.996	8.040	1.288	3.096	6.480	2.148	3.750	6.595
GPS	0.601	0.8573	1.720	0.582	0.777	2.334	0.634	0.863	2.157
C + G	0.752	0.828	1.747	0.572	0.635	2.332	0.908	0.832	1.896

The statistics results of different smoothing method are given in Tables 18.2, 18.3 and 18.4. The following conclusion can be got:

Table 18.4 Accuracy comparison of divergence free smoothed differential position (m)

	AB			AC			BC		
	E	N	U	E	N	U	E	N	U
COMPASS	0.617	1.443	3.021	0.462	1.042	2.008	0.543	1.182	2.222
GPS	0.232	0.342	0.730	0.209	0.310	0.569	0.245	0.330	0.697
C + G	0.245	0.338	0.697	0.217	0.275	0.571	0.249	0.288	0.622

1. When comparing Tables 18.2 and 18.1, we can find that the carrier smoothing can improve the differential position accuracy. Take baseline AB for example, the RMS of COMPASS in Up component reduced from 3.114 to 2.884 m.
2. Besides, after 100 s smoothing,the position accuracy of COMPASS/GPS is only a little better than that of GPS, the difference is a few centimeters. The possible reason is that the noise of GPS code measurement is reduced remarkablely, and the smoothing time for COMPASS is still too short, so it cannot make contribution to the position accuracy.
3. It can be seen from Table 18.3 that, when the ionosphere free smoothing are used, the position accuracy of both COMPASS and GPS are worse that of single frequency. The main reason is that the receiver noise and multipath error is amplified.
4. By comparing Tables 18.4 and 18.2, we can see that, the results of divergence free smoothing and single frequency smoothing are about the same when the smoothing time is 100 s.

Table 18.5 shows the variation of position accuracy of baseline AC (42.2 km) with different smoothing time. It can be seen that for the single frequency smoothing (SFCS), when the smoothing time increased, the position accuracy was improved firstly, when the smoothing time is larger than 900 s, the result is getting worse. While for the divergence free (DFree) smoothing, the position accuracy was improved significantly. The RMS of East, North, and Up component reduced from 0.462, 1.042 and 2.008 to 0.344, 0.404 and 1.259 m when the smoothing time increased form 100 to 3,600 s. Figure 18.4 shows the position error of DFree smoothing when the smoothing time is 100 and 3,600 s. Obviously, when the smoothing time is 3,600 s, the position error is much smaller.

Table 18.5 Variation of position accuracy with smoothing time (m)

Smoothing time (S)	SFCS			DFree		
	E	N	U	E	N	U
100	0.463	1.045	2.018	0.462	1.042	2.008
300	0.438	0.827	1.823	0.439	0.813	1.778
600	0.418	0.731	1.766	0.423	0.699	1.645
900	0.399	0.692	1.780	0.411	0.643	1.571
1,800	0.347	0.727	1.934	0.381	0.500	1.423
3,600	0.237	0.591	2.135	0.344	0.404	1.259

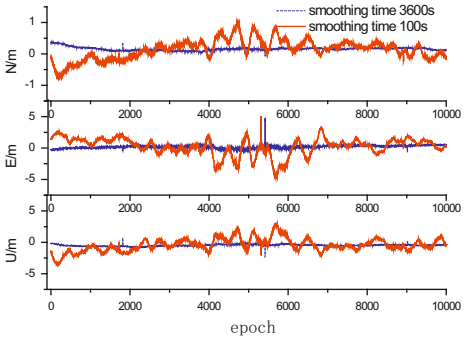


Fig. 18.4 Position error of different smoothing time

18.5 Conclusions

The establishment of COMPASS and COMPASS/GPS differential positioning is significant for enhancing the position accuracy. The position accuracy of COMPASS and COMPASS/GPS differential positioning is analyzed by using real data, and the following conclusion can be derived:

1. When no smoothing algorithm is used, the position accuracy of COMPASS single frequency differential positioning is about 2–4 m, and the accuracy of COMPASS/GPS is better than 1.5 m.
2. The position accuracy of COMPASS can be improved by carrier smoothing, when the smoothing time is 100 s, the position accuracy can be improved by about 20–50 cm. Longer smoothing interval can be used in application to further improve the differential position accuracy of COMPASS.
3. For the suggested 100 s smoothing time in GBAS, the difference between GPS and COMPASS/GPS is only a few centimeters,. Longer smoothing time can be used in application to further improve the differential position accuracy of COMPASS. The possible reason is the noise and multipath error of COMPASS code measurement is large, and the smoothing time is too short for COMPASS, and thus the improvement of position accuracy is not obvious.
4. The receiver noise and multipath error are amplified by ionosphere combination, so sometimes the position results of ionosphere free smoothing may be worse than that of single frequency. Therefore, the spatial gradient of ionosphere should be monitored to decide which smoothing algorithm should be used to get ideal position results.
5. The results of divergence free smoothing is about the same when the smoothing time is short. The advantage of divergence free smoothing is that it can improve the position accuracy of COMPASS by a longer smoothing time. This means a lot for COMPASS users.

Acknowledgments This work is supported by the National Natural Science Funds of China (Grant Nos. 41020144004; 41104022), the National "863 Program" of China (No: 2013AA122501), and the 2nd and 3rd China Satellite Navigation Conference (Grant Nos. CSNC-2011-QY-13; CSNC-2012-QY-03).

References

1. Miguel M, Maria D, Juan R, Manuel T (2011) GNSS, an evolving technology, current systems evolutions and future perspectives. ION GNSS 24th international technical meeting of the satellite division, 19–23 September 2011, Portland, QR. 3083-3101
2. Urlichich Y, Subbotin V, Stupak G, Dvorkin V, Povalyaev A, Karutin S (2011) GLONASS modernization. ION GNSS 24th international technical meeting of the satellite division, 19–23 September 2011, Portland, QR. 3125-3128
3. Yang Y, Li J, Xu J, Tang J, Guo H, He H (2011) Contribution of the compass satellite navigation system to global PNT users. Chin Sci Bull 56(26):2813–2819
4. Yang Y (2010) Progress, contribution and challenges of compass/beidou satellite navigation system. Acta Geodaetica et Cartographica Sinica 39(1):1–6
5. Konno H (2007) Design of an aircraft landing system using dual-frequency GNSS. Stanford University
6. Chen J (2001) Research of GPS ingerity augmentation. Institute of Surveying and mapping, Information Engineering University, Zhengzhou
7. Gan X (2008) Research on the integrity monitoring of GPS local area augmentation system. Harbin Engineer University, Harbin
8. Hideki Y, Tomoji T, Nobuaki K, Akio Y (2008) Evaluation of positioning accuracy with differential GPS/GLONASS. ION GNSS 21th international technical meeting of the satellite division, 16–19 Sept 2008, Savigation, GA, pp 2221–2226
9. Seo K Y, Park S H, Jang W S (2006) Performance Analysis of software-based GPS/GLONASS augmentation system for maritime DGNSS service. ION national technical meeting, 18–20 Jan 2006
10. Ward N. F (2006) Future of IALA DGNSS. ION national technical meeting, 18–20 Jan 2006
11. Shi C, Zhao Q, Hu Z, Liu J (2012) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. GPS solution, Published online: 19 May 2012. doi: 10.1007/s10291-012-0264-x
12. Hatch R (1982) The synergism of GPS code and carrier measurements. Proceedings of the 3rd international symposium on satellite doppler positioning, vol 2, pp 1213–1231
13. McGraw G (2006) How can dual frequency code and carrier measurements be optimally combined to enhance position solution accuracy. Inside GNSS 1(5):17–19

Chapter 19
A Method of Integrity Monitoring and Assessment for BeiDou Navigation Satellite System

Weiping Liu, Jinming Hao, Zhiwei Lv, Hongfei Xie and Yingguo Tian

Abstract The integrity of navigation satellite system is mentioned as a kind of ability that users can get warning in time when the system cannot be used. When the system cannot achieve the demand of integrity, its security will be declined. So the integrity, which is an important performance indicator, is always focused on by satellite navigation providers. In this paper, the concept of integrity is introduced. A method of integrity monitoring and assessment is presented. The real data is analyzed, and the method is validated. At last, the integrity of BeiDou is analyzed initially.

Keywords BeiDou navigation satellite system · Integrity · Monitoring and assessment · Stanford plot · Positioning precision · Performance indicator

19.1 Introduction

Following the "three-step" development strategy, BeiDou has announced Interface Control Document (ICD) and provided regional operational service to China and most parts of its surrounding areas since the end of 2012 [1, 2]. Based on the experience of GPS, GLONASS and Galileo, the performance of a navigation satellite system is related closely to availability, continuity and integrity, besides positioning precision. And Integrity is one of the most important performance indicators. For the development of BeiDou, we should focus on the research of integrity monitoring and assessment.

Integrity is defined as the ability that the system gives warning to users in time when the navigation service cannot be used. When the system cannot reach the demand of integrity, the security of service will decline, which may lead to heavy

W. Liu (✉) · J. Hao · Z. Lv · H. Xie · Y. Tian
College of Navigation and Aerospace Engineering, Information Engineering University, Zhengzhou, China
e-mail: lwpchxy@sina.comlwpchxy@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_19,

accident. So integrity, which is becoming more and more important, has attracted the system providers' and scientists' attention. In 1997, Colonel et al. [3] gave integrity requirements in GPS performance document. In 2004, Oehler et al. [4] introduced the concept of integrity about Galileo. In 2006, Sardón et al. [5] discussed the design of integrity processing unit on ground in Galileo. In 2008, integrity was included in GPS SPS PS [6]. And other people have also studied on integrity, such as Chen [7], Niu [8], and so on. The most typical method of integrity monitoring and assessment is Stanford Plot [9, 10]. The method, which can monitor and assess integrity very well in positioning area through the relationship among Protection Levels (PL), Alert Limits (AL) and Position Errors (PE), has been used in WAAS, EGNOS and Galileo.

Usually, Stanford Plot is used in static positioning mode. But if we can monitor and assess integrity of the system in both static and kinematic mode, the conclusion may be more believable and more comprehensive. So in this paper, a method of BeiDou kinematic positioning precision evaluation is progressed, and then the method, which is used for integrity monitoring and assessment of BeiDou in both mode, is present. At last, the initial analysis for integrity of BeiDou is made based on real data.

19.2 Method of Integrity Monitoring and Assessment

When integrity is assessed by Stanford Plot, positioning errors are very important. For static mode, positioning errors can be gotten by comparison with known coordinates. But for kinematic mode, positioning errors cannot be gotten so easily, so an method of BeiDou kinematic positioning precision evaluation is designed. And moreover, the method of Stanford Plot is also discussed in the section.

19.2.1 Precision Evaluation of Kinematic Positioning

The basic thought is: the highly precise GPS kinematic position is used as standard to assess BeiDou kinematic positioning precision. Multiple BeiDou receivers are considered to be tested so that the assessment results can be more believable. We design the platform for kinematic precision evaluation, as Fig. 19.1. The GPS precise position in A3 can be passed to other points (for example, A1) by the size and attitude of the platform. The GPS precise position can be gotten through precise relative positioning or precise point positioning. The size of the platform can be easily gotten. And the attitude of the platform can be gotten by inertial surveying system. To avoid error accumulation, we use IMU/GPS integrated inertial surveying system which is constituted by one IMU in C2 and two GPS receivers in C1 and C3, as Fig. 19.1. We will discuss how we pass the GPS precise position in A3 to other points here.

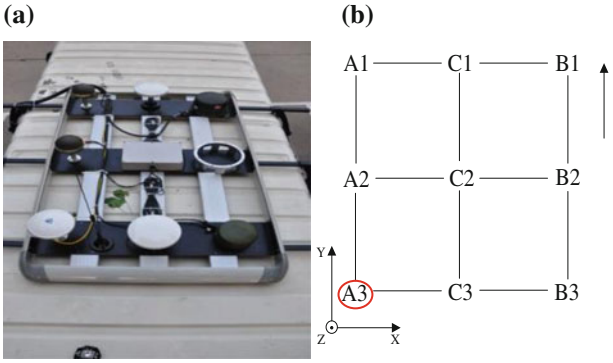


Fig. 19.1 Platform for kinematic precision evaluation. **a** Real photo. **b** Sketch map

At first, the body coordinate system O-XYZ is established. The origin is the phase center of the GPS receiver in A3. The axis of Y points to the direction of the velocity. The axis of Z points to the normal direction of the platform. O-XYZ is a right-hand coordinate system. In the following, the point A2 is taken as an example, and other points are similar.

As Fig. 19.1, the coordinate of $\overrightarrow{A_3A_2}$ is $(0, L_{23}, h_2 - h_3)$.L_{23} is the distance between A_2 and A_3, h_2, h_3 are the antenna heights of A_2 and A_3.

The relationship of body coordinate system and local coordinate system is shown as Fig. 19.2. In the figure, O-XYZ indicates body coordinate system, and O-NWU indicates local coordinate system (North–West-Up). They can be connected by $\psi/\theta/\gamma$.

We can get the angles of heading, roll and pitch from the IMU/GPS integrated inertial surveying system. And then, we have

$$\begin{cases} \psi = 270^0 - Heading \\ \theta = Pitch \\ \gamma = Roll \end{cases} \tag{19.1}$$

This leads to

$$\begin{aligned} \overrightarrow{A_3A_2}_{O-NWU} &= R_Z(-\psi)R_X(-\theta)R_Y(-\gamma)\overrightarrow{A_3A_2}_{\text{O\_XYZ}} \\ &= \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{pmatrix} \cdot \begin{pmatrix} 0 \\ L_{23} \\ h_2 - h_3 \end{pmatrix} \end{aligned} \tag{19.2}$$

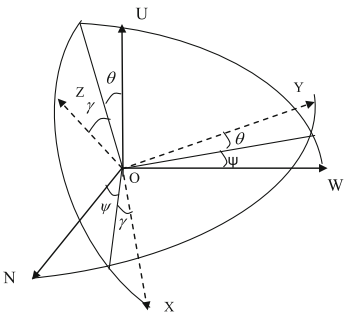


Fig. 19.2 Relation between body coordinate system and local coordinate system

And then, we can get the local coordinate of $\overrightarrow{A_3A_2}_{O-NWU}$, which can be written as $(N', W', U')^T$.

Assuming $\vec{\mathbf{r}}_{A_3}$ indicates GPS kinematic position in A_3, the earth fixed coordinate of A_2 can be expressed as

$$\vec{\mathbf{r}}_{A_2} = \vec{\mathbf{r}}_{A_3} + (ZG) \cdot \overrightarrow{A_3A_2}_{O-NWU} \tag{19.3}$$

$$(ZG) = R_z(-\pi - L)R_y(\pi/2 - B) \tag{19.4}$$

And then, the GPS precise position in A_3 can be passed to A_2. The transformation between WGS84 and CGCS2000 is also considered, and the positioning time is processed to keep synchronous. At last, the kinematic positioning precision of BeiDou can be assessed.

19.2.2 Method of Stanford Plot

Some concept is introduced at first. (1) Protection Levels (PL): the limit of position errors, which include Horizontal Protection Levels (HPL) and Vertical Protection Levels (VPL). (2) Alert Limits (AL): the warning limit of position errors, which include Horizontal Alert Limits (HAL) and Vertical Alert Limits (VAL). (3) Position Errors (PE): the differences between positioning results and the real positions, which include Horizontal Position Errors (HPE) and Vertical Position Errors (VPE).

Based on Stanford Plot, integrity can be assessed through the relationship among PL, AL and PE. PL can be calculated by the information from the navigation system [11, 12]. AL can be provided from the requirement of users, especially aerial users [13]. PE consists of static position errors and kinematic position errors. The former can be gotten through comparison with the known coordinates, and the latter can be gotten through the method in Sect. 19.2.1.

Stanford Plot usually includes horizontal plot and vertical plot. We take the vertical plot as an example (Fig. 19.3). And the relationship between integrity and PL/AL/PE is given as follows.

1. When $PE < PL < AL$, the system is available. $PL < AL$ indicates that the system is available by the system information. Based on this, $PE < PL$ indicates that position errors meet the limit of protection levels.
2. When $PE < AL < PL$ or $AL < PE < PL$, the system is unavailable. $AL < PL$ indicates that the system is unavailable. And meanwhile, $PE < PL$ indicates that position errors meet the limit of protection levels. In the above condition, the system is unavailable when $PE < AL$ or $AL < PE$.
3. When $PL < PE < AL$ or $AL < PL < PE$, the system may give Misleading Information (MI). When $PL < PE < AL$, $PL < AL$ indicates that the system is available, but $PL < PE$ indicates that position errors cannot meet the limit of protection levels. In other words, PL may not be believable. The information, which the system is available, may be Misleading Information. When $AL < PL < PE$, $AL < PL$ indicates that the system is unavailable, but $PL < PE$ indicates that position errors cannot meet the limit of protection levels. The information, which the system is unavailable, may be Misleading Information.
4. When $PL < AL < PE$, the system gives Hazardously Misleading Information (HMI). $PL < AL$ indicates that the system is available, but $AL < PE$ indicates that position errors are larger than alert limits. The information, which the system is available, is Hazardously Misleading Information.

Through the relationship above, we can monitor and assess integrity of the navigation system clearly by Stanford Plot.

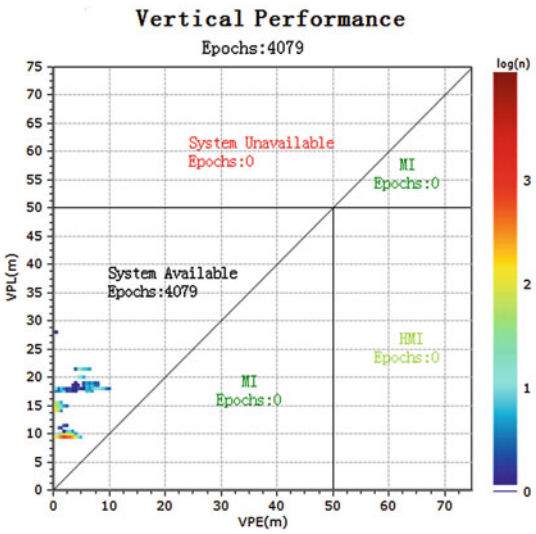


Fig. 19.3 Sketch map of stanford plot

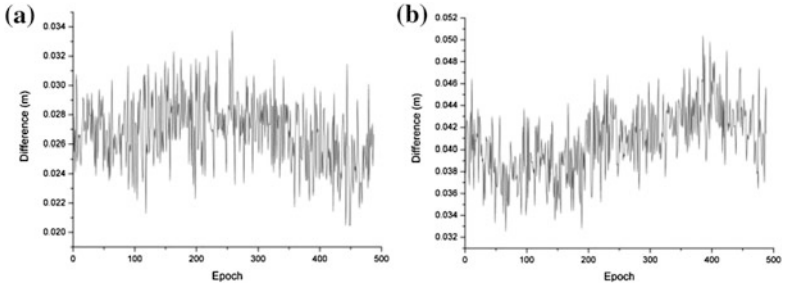


Fig. 19.4 Results of transformation. **a** A1. **b** B1

19.3 Numerical Examples and Analysis

19.3.1 Numerical Example 1

The example is designed to validate the rationality of the method in Sect. 19.2.1. Three GPS rover receivers are put on A3, A1 and B1, as Fig. 19.1. And the base station is in the radius of 20 km. Real-time GPS RTK positioning is carried out to the rover receivers. When the attitude of the platform can be gotten well, we pass the GPS precise position in A3 to A1 and B1 by the method in Sect. 19.2.1, and then compare them with the positioning results from GPS RTK. The differences are given in Fig. 19.4 and Table 19.1.

The results show that the differences are small between transformation results and GPS RTK positioning results in A1 and B1. RMS is better than 5 cm. Considering the precision of GPS RTK, the results are enough to validate the rationality of the method in Sect. 19.2.1.

19.3.2 Numerical Example 2

To validate the method of integrity monitoring and assessment in Sect. 19.2.2, a static experiment was processed in Zhengzhou on November 12th, 2011. B1C was used. The experiment span was 09:42–22:15, and the sample rate was 1 s. HPE and VPE were gotten from comparison with known coordinates. HPL and VPL were calculated from the system information. HAL and VAL were based on aerial

Table 19.1 Statistics of transformation results

	A1	B1
MAX (m)	0.034	0.050
MIN (m)	0.020	0.033
RMS (m)	0.027	0.041

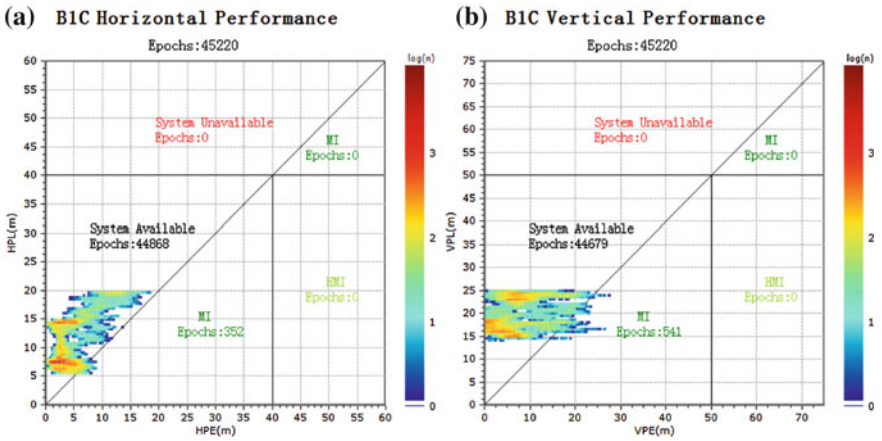


Fig. 19.5 Integrity performance in static experiment. **a** Horizontal performance. **b** Vertical performance

standard APV I (HAL = 40 m, VAL = 50 m). The results of the experiment are shown as Fig. 19.5. Table 19.2 contains the percents of all kinds of epochs.

In the long-time static experiment, numbers of epochs are zero when the system is unavailable or the system gives HMI. The percent of MI is 0.78 % in the horizontal direction and 1.20 % in the vertical direction. The system can provide normal service most of the time (horizontal: >99 %; vertical: ~99 %). In the experiment, the integrity of BeiDou is good, and it can be improved in the future.

19.3.3 Numerical Example 3

A kinematic experiment was also processed in Zhengzhou on August 11th, 2012. The velocity of the car was about 40 km/h. B1C was also used. The experiment span was 09:42–22:15, and the sample rate was 1 s. HPE and VPE were calculated by the method in Sect. 19.2.1. PL and AL were gotten by the same method with numerical example 2. We used multiple BeiDou receivers in the experiment. One of their results is given as Fig. 19.6 and Table 19.3.

In the kinematic experiment, vertical performance is comparable with the results in numerical example 2, but horizontal performance is worse. System

Table 19.2 Percents of epochs in static experiment

	Horizontal percent %	Vertical percent %
System unavailable	0	0
System available	99.22	98.80
MI	0.78	1.20
HMI	0	0

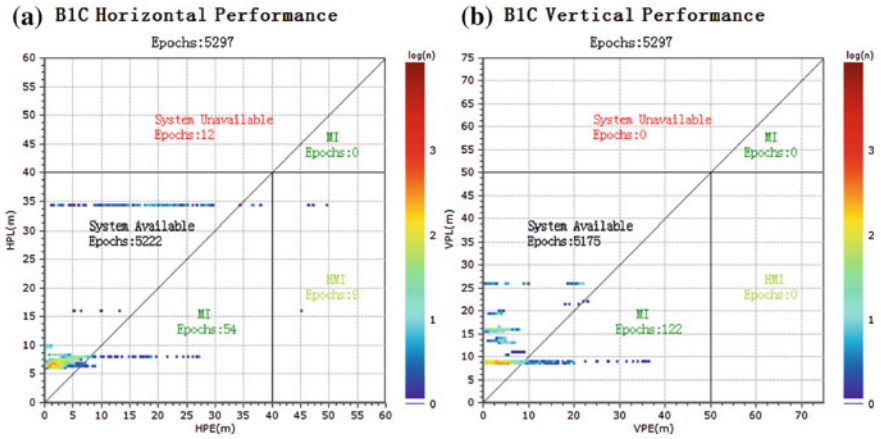


Fig. 19.6 Integrity performance in kinematic experiment. **a** Horizontal performance. **b** Vertical performance

Table 19.3 Percents of epochs in kinematic experiment

	Horizontal percent (%)	Vertical percent (%)
System unavailable	0.23	0
System available	98.58	97.70
MI	1.02	2.30
HMI	0.17	0

Unavailable indeed occurs in the horizontal direction and the percent reaches 0.23 %. After analysis, we find the reason is that the value of DOP decreased sharply in the span 16:46–16:58, which may be arose from "urban canyon".

19.4 Conclusion

The method of kinematic positioning precision evaluation for BeiDou is present in this paper. And then, we can use Stanford Plot to monitor and assess the integrity of BeiDou comprehensively in the static and kinematic mode. Based on real data, the method of BeiDou kinematic positioning precision evaluation is validated. And some experiments are processed to assess the integrity of BeiDou in the static and kinematic mode. We should emphasize that these experiments, which are only used to validate the rationality of the method in this paper, cannot represent the real performance of BeiDou enough. In the future, we will deal with more real data to analyse the integrity of BeiDou.

Acknowledgments Thanks to the members of our research group!

References

1. Yang Y (2010) Progress, contribution and challenges of Compass/Beidou satellite navigation system. Acta Geod et Cartographica Sinica 39(1):1–6
2. China satellite navigation office (2012) Report on the development of BeiDou navigation satellite system 2.0. CSNC
3. Colonel MA, Captain KJ (1997) Global position, velocity and timing determination capability capstone requirements document. In: Proceedings of the 10th international technical meeting of the satellite division of the institute of navigation (ION GPS 1997), pp 399–406
4. Oehler V, Luongo F, Boyero JP et al (2004) The Galileo integrity concept. In: Proceedings of the 17th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2004), pp 604–615
5. Sardón E, Mora E, Hernández C, Piedelobo JM (2006) Galileo integrity processing facility: preliminary design. In: Proceedings of the 19th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2006), pp 531–539
6. U.S. Department of defense (2008) Global position system standard positioning service performance standard, 4th edn. Washington, DC
7. Chen J (2001) Research of GPS integrity. PhD thesis, Information engineering university, Zhengzhou, China
8. Niu F (2009) Theory and technique on GNSS integrity augment. PhD thesis, Information engineering university, Zhengzhou, China
9. WADGPS laboratory (Stanford University) (1997) Precision approach metrics: accuracy, integrity, continuity and availability. http://waas.stanford.edu/metrics.htm
10. Tossaint M, Samson J, Toran F et al (2007) The Stanford-ESA integrity diagram: a new tool for the user domain SBAS integrity assessment. Navig-Los Angeles Wash 54(2):153–162
11. Lee Y (1995) New techniques relating fault detection and exclusion performance to GPS primary means integrity requirements. In: Proceedings of the 8th international technical meeting of the satellite division of the institute of navigation (ION GPS 1995), pp 1929–1939
12. Brown RG, Chin GY, Kraemer JH (1991) Update on GPS integrity requirements of the RTCA MOPS. In: Proceedings of the 4th international technical meeting of the satellite division of the institute of navigation (ION GPS 1991), pp 761–772
13. Feng S, Ochieng WY, Walsh D, Loannides R (2006) A measurement domain receiver autonomous integrity monitoring algorithm. GPS Solutions 10(2):85–96

Chapter 20
Research on Receiver Autonomous Integrity Monitoring Algorithm Using Genetic Algorithm Resampling Particle Filter

Pan He, Chun Tan and Huawen Huang

Abstract With the rapid development of global navigation satellite system, Receiver Autonomous Integrity Monitoring (RAIM) has attracted attention from many researchers and there still are a lot of unsolved problems in this field. Traditional RAIM algorithms are built upon Kalman filter under the assumption that satellite signal noise follows Gaussian distribution. However, with the impact of ionospheric delay error and other factors, the measurement noise usually doesn't follow Gaussian distribution. Under strong interference and harsh environmental conditions, particle filter is often employed to improve the efficiency of RAIM with non-Gaussian distribution errors. But those algorithms on particle filter are poor in convergence accuracy and stability because of particle degeneracy and this paper aims to provide a solution for particle degeneracy. Using the idea of approximate probability, this paper also takes use of particle filter in RAIM for fault detection and employs the idea of genetic operations to avoid particle degrading too early. Based on the classic particle filter procedure, simulation binary crossover operator, polynomial mutation operator and roulette wheeling selection method are used to produce new generation of particles from old generation in resampling process to increase particle diversity in state space as well as keeping good performance particles. The modified particle filter is then used in the fault detection and isolation process of RAIM with cumulative log likelihood ratio test. Finally, experiment are conducted using IGS tracking station observation data and the results showed that, our algorithm could be used to detect and isolate faults under non-Gaussian noise environment, which proved the efficiency of our algorithm in RAIM. Comparing

P. He (✉) · C. Tan · H. Huang
Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China
e-mail: hepan@cigit.ac.cn

C. Tan
e-mail: tanchun@cigit.ac.cn

H. Huang
e-mail: hhw@cigit.ac.cn

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_20,

with RAIM algorithms based on traditional particle filters, our algorithm could improve the fault detection accuracy and convergence rate under non-Gaussian noise conditions as well as avoiding particle degeneracy.

Keywords Global navigation satellite system · Receiver autonomous integrity monitoring · Particle filter · Genetic algorithms

20.1 Introduction

When the error of a satellite navigation system exceeds allowable limits and causes the system incapable of providing required functions, integrity refers to the ability to issue an alarm in time, to notify corresponding users and to terminate signal processing. The integrity monitoring indicators could be used to evaluate the reliability of navigation systems [1]. Receiver Autonomous Integrity Monitoring (RAIM) takes use of the redundant information from receiver itself to detect and isolate faults of satellite signals, so it is the most popular method for integrity monitoring [2]. While Kalman filter demonstrates better monitoring performance than traditional methods in RAIM [3, 4], it asks the noise of satellite signals to follow Gaussian distribution, which restricts the application of Kalman filter in non-Gaussian distribution error detection and isolation situations [5].

Aiming at dealing with non-Gaussian distribution measurement noise, researchers propose to employ non-linear filter methods. To provide a high integrity GPS/INS position system, Ref. [6] takes use of Extended Kalman filter to implement position algorithm and monitor integrity. Reference [7] and [8] utilize MCMC particle filter and auxiliary particle filter with log likelihood ratio test to implement RAIM algorithm. With the increase of sample particles size, particle filter let the distribution of particles gradually move towards the posterior probability density of the state. It exhibits obvious advantages in fault detection and isolation with non-Gaussian distribution noise [9, 10]. Yet, in existing particle filter methods used in RAIM, particles often degrade too early and the algorithms may have a bad performance on convergence accuracy or stability. To avoid particle degeneracy, Ref. [11] summarizes that various methods like auxiliary sampling-resampling, regularized sampling could be used, yet the efficiency of the above methods can further be improved. Genetic operators could produce a new generation of particles, which may increase the diversity of particles in state space and eliminate the disadvantage of random resampling [12].

To avoid particles degeneracy, this paper adopts genetic algorithms in particle resampling process. New particles are generated by crossover and mutation operators, used to replace old ones under certain circumstance and to form a new group of particles. The modified particle filter algorithm based on genetic-resampling is then used in RAIM to perform fault detection and isolation. Finally, experiments are conducted to prove the efficiency and advantages of our algorithm in RAIM.

20.2 Particle Filter Based on Genetic Algorithm Resampling

An example of dynamic state space model at time k is shown by Eq. (20.1). x_k and z_k are state vector and measurement vector, f and h are state transition function and measurement function. w_k and v_k are the state and measurement noise.

$$\begin{aligned} x_{k+1} &= f(x_k, w_k) \\ z_k &= h(x_k, v_k) \end{aligned} \tag{20.1}$$

Recursive Bayesian estimation takes advantage of mathematical models and observation data to estimate the probability density function of unknown targets recursively. Particle filter is a Monte Carlo implementing method of recursive Bayesian estimation, and its core idea is to use a set of weighted random particles to approximately present the unknown posterior probability density $p(x_k|z_{1:k})$, shown by Eq. (20.2), where $z_{1:k}$ is the set of measurement vector $\{z_1, \ldots, z_k\}$. N is the size of the sample set, ω_k^i is the weight of each particle at time k, δ is Dirac pulse function, and $\tilde{x}_k^i$ is the ith random sample for state vector x_k.

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i \delta\left(x_k - \tilde{x}_k^i\right), \text{where } \sum_{i=1}^{N} \omega_k^i = 1 \tag{20.2}$$

Adopting the priori probability density $p\left(\tilde{x}_k^i|\tilde{x}_{k-1}^i\right)$ as the importance sampling function, the recurrence formula for weight is shown by Eq. (20.3).

$$\omega_k^i = \omega_{k-1}^i \frac{p\left(z_k|\tilde{x}_k^i\right)p\left(\tilde{x}_k^i|\tilde{x}_{k-1}^i\right)}{q\left(\tilde{x}_k^i|\tilde{x}_{0:k-1}^i, z_{1:k}\right)} = \omega_{k-1}^i p\left(z_k|\tilde{x}_k^i\right) \tag{20.3}$$

Higher value of ω_k^i indicates the particle $\tilde{x}_k^i$ more close to the real state x_k. So in sampling importance resampling method, ω_k^i is often used as the main indicator to select particles. It helps to keep better performance particle but decreases the diversity of particles in state space, which may lead to particle degeneracy. In order to increase particle diversity as well as keeping better particles, in our algorithm, simulation binary crossover operator is used in resampling to produce new particles $\tilde{x}_k'^i$ and $\tilde{x}_k'^j$ from old ones, shown by Eq. (20.4). Polynomial mutation operator is also used to generate new particle from old one as Eq. (20.5). η and η_m denote the predefined crossover and mutation index.

$$\begin{aligned} \tilde{x}_k'^i &= 0.5((1-\beta)\tilde{x}_k^i + (1+\beta)\tilde{x}_k^j) \\ \tilde{x}_k'^j &= 0.5((1+\beta)\tilde{x}_k^i + (1-\beta)\tilde{x}_k^j) \end{aligned}, \beta = \begin{cases} (2\alpha)^{\frac{1}{\eta+1}}, 0 \le \alpha < 0.5 \\ 1/((2(1-\alpha))^{\frac{1}{\eta+1}}), 0.5 \le \alpha < 1 \end{cases}, \alpha \sim N[0,1] \tag{20.4}$$

$$\tilde{x}_k^{\prime i} = \begin{cases} \tilde{x}_k^{\prime i} + (2r)^{\frac{1}{\eta_m+1}} - 1, 0 \le r < 0.5 \\ \tilde{x}_k^{\prime i} + 1 - [2(1-r)]^{\frac{1}{\eta_m+1}}, 0.5 \le r < 1 \end{cases}, r \sim N[0,1] \tag{20.5}$$

While the new particles are generated from crossover and mutation operators, $eval(\tilde{x}_k^i) = p(z_k|\tilde{x}_k^i)$ is used as fitness function. The roulette wheeling selection method is used to select new generation of particles from both old and new ones, according to their fitness value. The selection probability for each particle is defined as p_k^i, by Eq. (20.6). Using roulette wheeling selection method, the particle with higher fitness value has a higher probability to be selected.

$$p_k^i = \begin{cases} p_k^{i-1} + eval(\tilde{x}_k^i) \Big/ \sum_{j=1}^{n} eval(\tilde{x}_k^j), & 1 \le i \le N \\ 0, \ i = 0 \end{cases} \tag{20.6}$$

20.3 RAIM Algorithm Based on Genetic-Resampling Particle Filter

In this section, the genetic algorithm resampling particle filter described above is used in the RAIM algorithm to detect and isolate faults in satellite signal data.

Considering the static positioning scenario, the receiver state model and measurement model for receiver positioning are listed as Eq. (20.7) [13].

$$\begin{aligned} \mathbf{x}_k &= \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k \\ \rho_k &= h_k(\mathbf{x}_k) + v_k \end{aligned} \tag{20.7}$$

$\mathbf{x}_k = [r_x, r_y, r_z, \Delta\delta]_k^T$ is the state vector of the receiver, including X, Y, Z coordinates and clock offset of receiver at time k. ρ_k is the measurement vector (pseudorange). $\mathbf{F}_k$ is an identity matrix of size 4×4. $\mathbf{w}_k$ and v_k are system state noise and measurement noise. h_k represents the pseudorange calculation function: $\sqrt{(r_x - x_0)^2 + (r_y - y_0)^2 + (r_z - z_0)^2} - c\Delta\delta$. (x_0, y_0, z_0) is the three-dimensional coordinates of the current satellite position, and c is the light speed.

20.3.1 Fault Detection and Isolation Process

If a receiver could receive signals from m satellite and the measurement pseudorange value of jth satellite at time k is ρ_k^j, then $m + 1$ observation subset is set up as $y_k^0 = \left[\rho_k^1, \rho_k^2, \rho_k^3, \ldots, \rho_k^m\right]^T$, $y_k^j = \left[\rho_k^1, \rho_k^2, \ldots, \rho_k^{j-1}, \rho_k^{j+1}, \ldots, \rho_k^m\right]^T, j = 1, 2 \ldots m.$

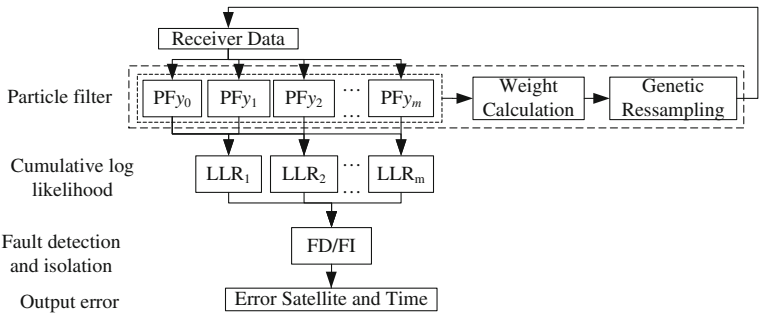


Fig. 20.1 Fault detection and isolation process based on particle filter and likelihood evaluation

Based on the above $m + 1$ observation subset, satellite faults detection and isolation process is shown by Fig. 20.1. $m + 1$ particle filters are used to predict, calculate cumulative log likelihood value, and capture the satellite faults using consistency test. Meanwhile, weights and genetic algorithm are used in the resampling process to resample and adjust the particles in case the particles degrade too early.

20.3.2 Improved RAIM Based on Genetic-Resampling Particle Filter

Based on the above fault detection and isolation process and genetic resampling particle filter, the framework of RAIM algorithm is listed below.

1. Initialization: Randomly generate N initial particles $\mathbf{x}_0^i \sim p(\mathbf{x}_0), i = 1, 2, \ldots N$ according to the approximate initial value of state vector and the priori probability distribution $p(\mathbf{x}_0)$. Generate the initial weights of particles $\omega_0^i = 1/N, i = 1, 2, \ldots N$.
2. Prediction: Generate $\mathbf{x}_k^i$ from $\mathbf{x}_{k-1}^i$ according to Eq. (20.7). For the m satellites, the three-dimensional coordinates for the jth satellite is $\left(x_0^{(j)}, y_0^{(j)}, z_0^{(j)}\right)$, $j = 1, 2 \ldots m$. Based on the value of $\mathbf{x}_k^i$ and the measurement function h_k, the estimate pseudorange for each satellite is evaluated as $\tilde{\rho}_k^{i(j)} = h_k^{(j)}\left(\mathbf{x}_k^i\right)$.
3. State likelihood evaluation: Upon receiving the real pseudorange for each satellite $\rho_k^{(j)}$, the likelihood of the ith particle at time k (denoted by $l_k^{i(j)}$) is evaluated as Eq. (20.8), where c_k denotes the variance of v_k. Assuming that the observation data for each satellite is statistically independent, the likelihood of each particle for each observation subset $p\left(y_k^j | \mathbf{x}_k^i\right)$ is calculated as the product of likelihood value of each particle for each satellite. The likelihood of the whole

set of particles for each observation subset at time k: $p\left(y_k^j|\mathbf{x}_k\right)$ is evaluated as the mean value of each particle according to Eq. (20.9).

$$l_k^{i(j)} = p\left(\rho_k^{(j)}|\mathbf{x}_k^{i(j)}\right) = \frac{1}{\sqrt{(2\pi)c_k}} e^{-\frac{\left(\rho_k^{(j)} - \tilde{\rho}_k^{i(j)}\right)^2}{2c_k}} \tag{20.8}$$

$$p\left(y_k^j|\mathbf{x}_k\right) \approx \frac{1}{N}\sum_{i=1}^{N} p\left(y_k^j|\mathbf{x}_k^i\right) = \frac{1}{N}\sum_{i=1}^{N}\prod_{w=1\&w\neq j}^{m} l_k^{i(w)}, j = 0, \ldots, m \tag{20.9}$$

4. Cumulative log likelihood ratio (LLR) evaluation: From time a to time b, the cumulative LLR is calculated as Eq. (20.10).

$$S_a^{b(j)} = \sum_{k=a}^{b} \ln\frac{p\left(y_k^j|\mathbf{x}_k\right)}{p\left(y_k^0|\mathbf{x}_k\right)}, 1 \le j \le m \tag{20.10}$$

5. Fault detection and isolation: Fault detection function β_k is presented by the maximum cumulative LLR among each satellite in x period of time, as Eq. (20.11), in which x is the accumulation time window size.

$$\beta_k = \max_{1\le j\le m}\left\{\max_{k-x+1\le a\le k}\{S_a^{k(j)}\}\right\} \tag{20.11}$$

At any time k, if $\beta_k > \tau$ stands (τ is the predefined detection threshold value), there are errors in the observation data and the algorithm will calculate the set of indexes of fault data $g = \arg\{\max\{S_{k_0}^{k(j)}\}\}$, isolate this part of data in position calculation and return to Step 2. k_0 represents the time to issue an alarm. If $\beta_k \le \tau$ stands, no faults have occurred and the algorithm directly jumps to Step 6.

6. Update: The weight of each particle $\tilde{\omega}_k^i$ is calculated as Eq. (20.12).

$$\tilde{\omega}_k^i = \omega_k^i \Big/ \sum_{i=1}^{N}\omega_k^i, \omega_k^i = \omega_{k-1}^i p\left(y_k^j|\mathbf{x}_k^i\right) \tag{20.12}$$

Genetic resampling method is conducted using particle weights as follows:

(1) Crossover: Randomly select $N \times p_c$ pairs of particles using crossover rate p_c and generate two new particles from each pair of particles using Eq. (20.4).
(2) Mutation: Randomly select $N \times p_m$ particles according to mutation rate p_m and generate new particles from each particle using Eq. (20.5).
(3) Selection: Take $\tilde{\omega}_k^i$ (the normalized value of $p\left(z_k|\tilde{x}_k^i\right)$) as the fitness value of each particle, and use roulette wheeling method to select new generation of

particles from old generation using Eq. (20.6). After that, set the weight of each particle as $\omega_k^i = 1/N, i = 1, 2, \ldots, N$ and return to Step 2.

20.4 Empirical Studies and Result Analysis

To evaluate and compare the efficiency of our algorithm with other RAIM algorithms, IGS tracking station data provided by University of California, San Diego (UCSD) [14] is used in this section, including the precise ephemeris data for satellite position coordinates calculation and the observation data for measurement pseudorange calculation. The time range is from 10:00:00 am to 13:00:00 pm on October 26th, 2012. The time interval between data acquisition is 30 s and there are totally 360 groups of observation data. During this period, the minimal number of visible satellites is 6 and the corresponding satellites ID are: G21, G18, G3, G30, G6 and G16. After formal evaluation we know that during this period of time, all satellites are working in a normal state.

20.4.1 Test for Failure-Free Case

Following parameter settings used in existing experiments [8], we first conduct RAIM in failure-free case. The number of particles N is 1000, time window size x of fault detection is 50. To achieve balance between particle diversity and convergence rate, after several times of test, crossover rate p_c and mutation rate p_m are set as 0.8 and 0.1. Sampling importance resampling particle filter (SIR-PF) is used to compare with our genetic-resampling particle filter (Genetic-PF). The cumulative LLR for each subset of observation data using two algorithms is shown in Fig. 20.2 along with the change pattern of fault detection function β_k. Comparing to normal GPS receiver observation data, the sampling rate of IGS station data is relatively low (1/30 s), so it is difficult to estimate the measurement error distribution in observation data. According to Fig. 20.2, cumulative LLR got by SIR-PF couldn't converge to a stable value because of particle degeneracy. On the contrary, cumulative LLR got by Genetic-PF fluctuated but converged below 300. The fault detection function showed similar results. The particle degeneracy trend of SIR-PF is also shown in Fig. 20.2d. While the size of unique particles stayed around 600 for Genetic-PF, the according size decreased gradually using SIR-PF.

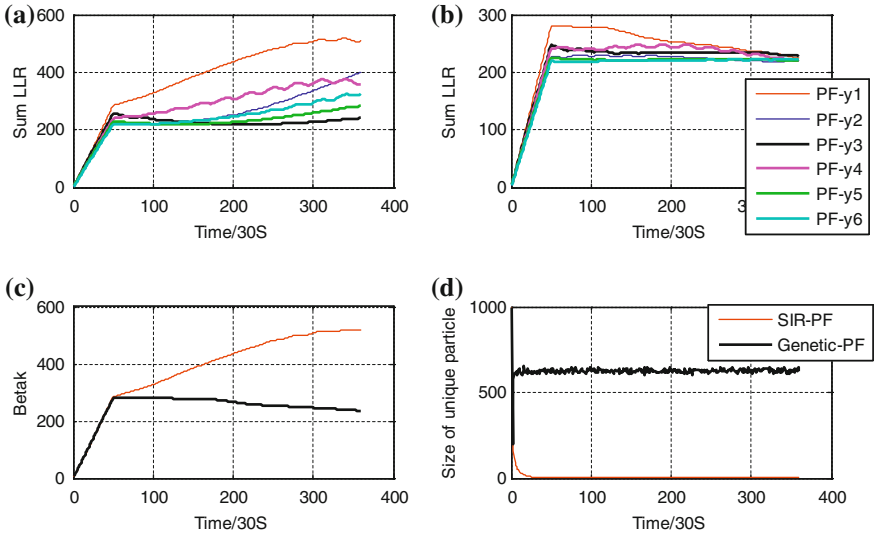


Fig. 20.2 Comparison of SIR-PF and Genetic-PF in failure-free case. **a** Cumulative LLR using SIR-PF. **b** Cumulative LLR using Genetic-PF. **c** Fault detection function value. **d** Size of unique particles in state space

20.4.2 Test for Fault-Injection Case

To test the efficiency of our algorithm when fault occurs, we deliberately inject errors in the 200th to 360th observation data of the 4th satellite. Two kinds of errors are injected, including bias error and ramp error. Because the sampling rate is relatively low and the inherent error in the data is relatively large, the range of injected bias error is chosen as 30 $\sim$ 100 m. Ramp errors start at 200th epoch with three different gradient values, 0.2, 0.3, 0.5 m/30 s, respectively. The cumulative LLR for each subset of observation data using Genetic-PF is shown in Fig. 20.3. When fault occurred, there wold be obvious change in the value of cumulative LLR, especially for the result of particle filter using observation subset y_4, which suggested that there were errors in the observation data of the 4th satellite, and this part of data should be excluded from the observation data. As the fault value increased, the change of cumulative LLR value also became bigger. Fig. 20.4 illustrates the fault detection situation of both filters on different fault-injection scenario. When the injected error was relatively large (100 m), both methods could detect the error easily through the change of β_k. However, when the error was relatively small (50 m), it was difficult to capture error using β_k obtained by SIR-PF while Genetic-PF had a much better performance. Moreover, small errors (30 m) couldn't be detected by SIR-PF entirely and Genetic-PF provided the possibility to detect small errors. The ramp error injection experiments showed the similar results.

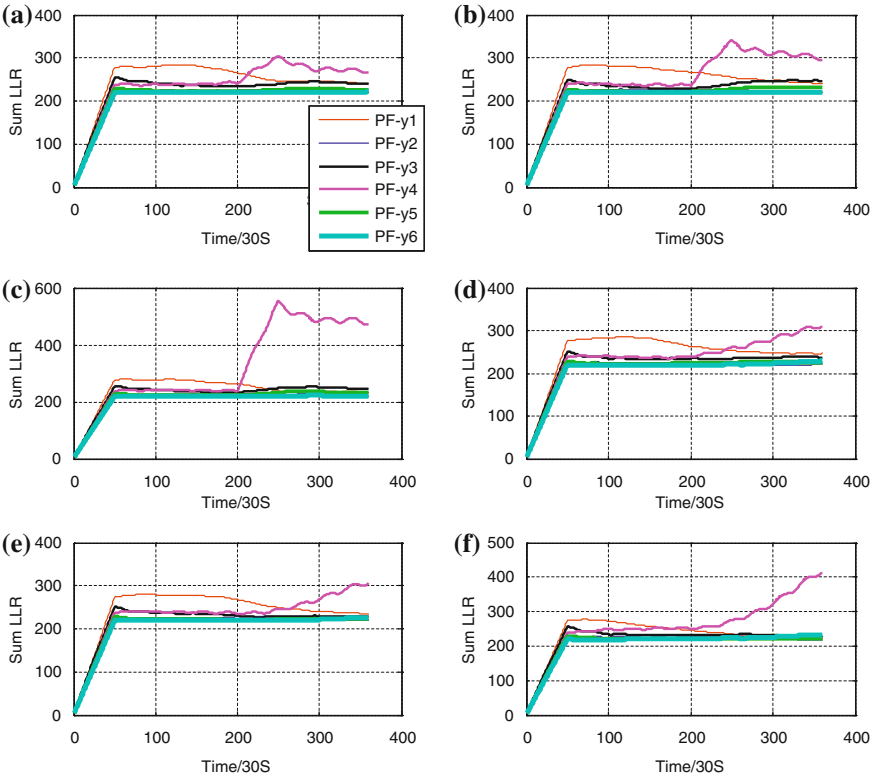


Fig. 20.3 Cumulative log likelihood ratio using Genetic-PF after fault injection. **a** Bias error 30 m. **b** Bias error 50 m. **c** Bias error 100 m. **d** Ramp error 0.2 m/30 s. **e** Ramp error 0.3 m/30 s. **f** Ramp error 0.5 m/30 s

20.5 Conclusion and Future Work

Aiming at tackling with non-Gaussian distribution noise, non-linear filter method is used to detect faults in current RAIM algorithms. In this paper, genetic resampling method is taken use of to avoid particle degeneracy and to improve the estimation accuracy of receiver observation data. Observation data from GPS observation station is employed to test the algorithm and experimental results showed that: genetic algorithm resampling particle filter applied to the fault detection and isolation problem when it was difficult to estimate the distribution of error in observation data; under both failure-free and non-Gaussian distribution noise circumstance, our algorithm outperformed traditional particle filter in both converge rate and fault detection accuracy.

Although genetic resampling particle filter outperformed existing algorithms in detection accuracy, the complexity of the algorithm has increased and genetic

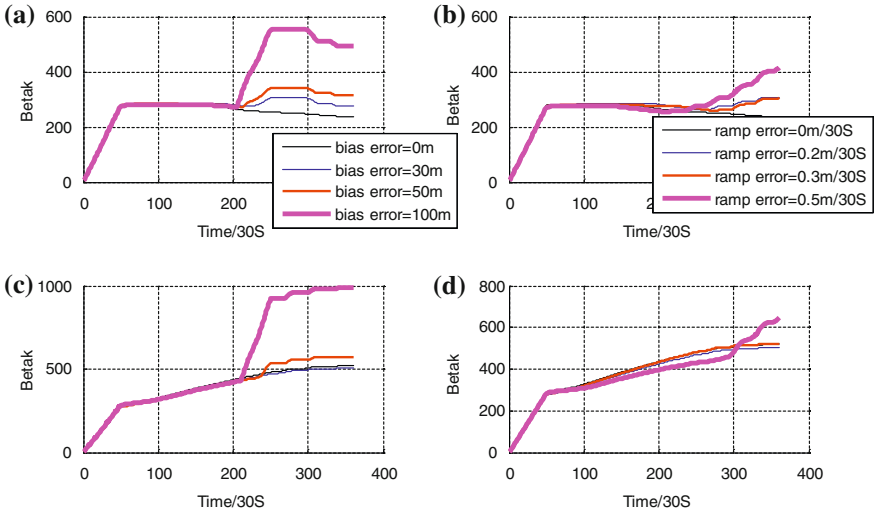


Fig. 20.4 Fault detection function value using different particle filters after fault injection. **a** Genetic-PF on bias error. **b** Genetic-PF on ramp error. **c** SIR-PF on bias error. **d** SIR-PF on ramp error

operations further increase the time complexity, which impact the overall performance of the RAIM algorithm. So, in the next step of work, we aim to improve the algorithm performance and build a fast fault detection method for RAIM.

Acknowledgments This work is supported by the Major Project of the Key Technologies R&D Program of CQ CSTC (No.CSTC2011GGC40008).

References

1. Zhang J (2011) New development and application of airspace surveillance technology. Acta Aeronaut et Astronaut Sin 32(1):1–14
2. Hewitson S, Wang JL (2006) GNSS receiver autonomous integrity monitoring (RAIM) performance analysis. GPS Solutions 10(3):155–170. doi:10.1007/s10291-005-0016-2
3. Hewitson S, Wang JL (2007) GNSS receiver autonomous integrity monitoring with a dynamic model. J Navig 60(2):247–263. doi:10.1017/so373463307004134
4. Hu GR, Ou JK (1999) The improved method of adaptive Kalman filtering for GPS high kinematic positioning. Acta Geod et Cartographica Sinica 28(4):290–294
5. Xu XH, Yang CS, Liu RH (2012) Review and prospect of GNSS receiver autonomous integrity motoring. Acta Aeronaut et Astronaut Sin, 33(x), published online
6. Toledo-Moreo R, Zamora-Izquierdo MA, Ubeda-Minarro B, Gomez-Skarmeta AF (2007) High-integrity IMM-EKF-based road vehicle navigation with low-cost GPS/SBAS/INS. IEEE Trans Intell Transp Syst 8(3):491–511. doi:10.1109/tits.2007.902642
7. Wang ES, Zhang SF, Hu Q (2009) Research on GPS receiver autonomous integrity monitoring algorithm based on MCMC particle filtering. Chin J Sci Instrum 30(10):2208–2212

8. Ahn J, Rosihan Won DH, Lee YJ, Nam GW, Heo MB, Sung S (2011) GPS integrity monitoring method using auxiliary nonlinear filters with log likelihood ratio test approach. J Electr Eng Technol 6(4):563–572. doi:10.5370/jeet.2011.6.4.563
9. Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 50(2):174–188. doi:10.1109/78.978374
10. Li P, Kadirkamanathan V (2001) Particle filtering based likelihood ratio approach to fault diagnosis in nonlinear stochastic systems. IEEE Trans Syst Man Cybern Part C-Appl Rev 31(3):337–343
11. Hu SQ, Jing ZL (2005) Overview of particle filter algorithm. Control and Decis 20(4):361–365 371
12. Lin K, Shi JH, Deng ZM (2012) Multipath target tracking based on genetic algorithm particle filter. Electron Meas Technol 35(2):50–54
13. Da R, Lin CF (1995) A new failure detection approach and its application to GPS autonomous integrity monitoring. IEEE Trans Aerosp Electron Syst 31(1):499–506
14. SOPAC. http://sopac.ucsd.edu/cgi-bin/dbDataByDate.cgi. Cited 19 Jan 2013

Chapter 21
Real-Time Monitoring of Regional Ionosphere Based on Polynomial Model with Multi-Station

Wenfeng Nie, Wusheng Hu, Zhiyue Yan and Shuguo Pan

Abstract The purpose of this paper is to monitor the ionospheric total electron content (TEC) in real time. Firstly, the receiver differential code bias (DCB) and the satellite DCB are calculated by assuming that they are constant value during the whole day. Comparing with the satellite DCB given by the Center for Orbit Determination in Europe (CODE), this paper also analyzes the DCB stability of receiver and satellite. And secondly, according to the stability of DCB, we calibrate vertical TEC (VTEC) by using DCB of the former day, and establish the polynomial model with multi-station in single epoch to monitor the changes of ionosphere. The result of all epoches within a day shows that the average internal RMS of this method can be better than 1 TECU (10^{16} el/m^2), and the average external RMS can be about 1 TECU.

Keywords VTEC · Real-time monitor · DCB

21.1 Introduction

As a most important part of space environment, the ionosphere affects the electromagnetic waves and human activities. The ionosphere research is both complicated and meaningful. And it is not only benefited to get knowledge of the ionosphere itself, seek ways to overcome damages by ionosphere, pursue methods to serve mankind, but also to motivate the research and development of the ionosphere, which is related to the earth science field [1]. With the appearing of the

W. Nie (✉) · W. Hu · Z. Yan
School of Transportation of Southeast University, Nanjing, China
e-mail: wenfengnie@126.com

S. Pan
School of Instrument Science and Engineering, Nanjing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_21,

global positioning system (GPS), the monitor of the ionosphere by GPS receives widespread attention, because of the advantages of GPS: high accuracy, high reliability, huge amount of data, 24-h observation and so on [2].

In the research of ionospheric monitor, many researchers put forward different indexes to judge the level of ionospheric irregularities and scintillation. Li et al. [3] and Chen et al. [4] introduced the phase scintillation index and the amplitude scintillation index S4 that based on the power of the GPS signal. Lambert et al. [5] analyzed the ionospheric status in middle latitude of Europe region and then promoted the ionospheric I95 index which was based on the double differential residual of ionosphere. In 1996, Aarons et al. [6] put forward ROT index which means the rate of TEC, to indicate the irregularities of ionosphere. Furthermore, ROTI index, which means the standard deviation of ROT was proposed in 1997 [7]. In summarize, these indexes can't provide the change of ionosphere in real-time according to the quantities of TEC. Aimed at the drawbacks mentioned above, Geng et al. [8] established polynomial model using single station in the regional continuous operation reference station (CORS) to monitor the change of TEC in real-time. The results showed that the internal RMS and external RMS of the model are separately about 3 TECU and 4 TECU. The accuracy of the model is quite low because of the small number of the ionosphere pierce point when using only one station.

Based on the conclusions of Geng et al. [8], this paper establishes polynomial model with multi-station to monitor the change of ionosphere in real-time. Firstly, the receiver DCB and the satellite DCB are calculated by assuming that they are both constant value during the whole day. Comparing with the satellite DCB given by CODE, this paper also analyzes the DCB stability of receiver and satellite. And secondly, according to the stability of DCB, we calibrate VTEC by using DCB of the former day, and establish the polynomial model with multi-station in single epoch to monitor the changes of ionosphere and also the accuracy of the model is analyzed.

21.2 Principal of Calculating TEC Using GPS

Mathematically, the code observations, denoted as P_1, P_2, can be described as [9]

$$\left\{ \begin{array}{l} P_{1j}^{i} = \rho_{j}^{i} + c(dt^{i} - dt_{j}) + d_{tropj}^{i} + d_{ion1j}^{i} + dq_{1j} + dq_{1}^{i} \\ P_{2j}^{i} = \rho_{j}^{i} + c(dt^{i} - dt_{j}) + d_{tropj}^{i} + d_{ion2j}^{i} + dq_{2j} + dq_{2}^{i} \end{array} \right\} \tag{21.1}$$

where

ρ is the true geometric range between receiver and satellite (m)
c is the speed of light (m/s);
dt^i is the satellite clock error with respect to GPS time (s);
dt_j is the receiver clock error with respect to GPS time (s);
d_{trop} is the tropospheric delay (m);

d_{ion} is the ionospheric delay (m);
dq^i is the satellite DCB (m);
dq_j is the receiver DCB (m);

Meanwhile, if the influence of the higher order term of the ionospheric delay is ignored, the expression of ionospheric delay can be shown as follows:

$$d_{ion} = 40.28 \frac{TEC}{f^2} \tag{21.2}$$

where

TEC is the total electron content in the path of signal pass through (TECU) f is the frequency of L1 or L2, f_{ 1} = 1575.42 MHz, f_{ 2} = 1227.60 MHz

Combined the Eqs. (21.1 and 21.2),

$$P^i_{kj} = \Psi + d_{ion} + dq_{kj} + dq^i_k \tag{21.3}$$

where

$$\Psi = \rho^i_j + c\,(dt^i - dt_j) + d^i_{tropj}$$

So, for a dual-frequency GPS receiver, the TEC can be calculated by the equation in the following:

$$P^i_{2j} - P^i_{1j} = 40.28 \frac{TEC}{f_2^2} - 40.28 \frac{TEC}{f_1^2} + dq_j + dq^i \tag{21.4}$$

$$TEC = 9.52437\,(P^i_{2j} - P^i_{1j} - dq_j - dq^i) \tag{21.5}$$

where $dq_j = dq_{2j} - dq_{1j}$, $dq^i = dq^i_2 - dq^i_1$, separately as the receiver's DCB and the satellite's DCB.

According to Eq. (21.5), we find that the key to calculate TEC is to improve the accuracy of the code observation and to detach DCB of the receiver and satellite. For the reason that code observation is not stable when satellite rises up or drops down, flags are set both after rise and before down in ten epoch to exclude calculating. As to the detach of DCB, the following will show the method by assuming they are constant value in the whole day.

21.3 Calculating DCB and Its Stability

According to Zhang [10], the value of DCB can be more stable and consist by assuming they are constant in the whole day when calculated. So the method is accepted in this paper.

Three-dimensional (latitude, longitude and time) Taylor series expansions are used to locally represent the VTEC:

$$VTEC = \sum_{i=0}^{n} \sum_{k=0}^{m} E_{ik} (\varphi - \varphi_0)^i (S - S_0)^k \tag{21.6}$$

where

(φ_0, λ_0)	is the latitude and longitude of the central point of the region;
(φ, λ)	is the latitudes and longitudes of the ionosphere pierce points;
S_0	is the hour angles of the Sun on the middle epoch of the observation session observed at the central point, e.g. $S - S_0 = (\lambda - \lambda_0) + (t - t_0)$, t is the observation epoch, t_0 is the middle eoch;
(n, m)	are the maximum orders of the expansion in latitude and longitude;
E_{ik}	are the unknown coefficients of the Taylor series expansion;

The conversion from vertical TEC into slant TEC is written as:

$$TEC = mf(z) \cdot VTEC \quad mf(z) = 1/\cos(z') \tag{21.7}$$

with $\sin z' = R \cdot \sin z/(R_0 + H)$ where

$mf\ (z)$	denotes the single-layer mapping function;
z, z'	are the (geocentric) zenith distances at the height of the station and the single-layer;
R	is the radius with respect to the station considered;
$R_0 = 6371$ km	is the mean radius of the Earth, and

H is the height of the single-layer above the Earth's mean surface. The height H of this idealized layer is usually set to the height of the expected maximum electron density: e.g. H = 450 km.

More details about the expansion and the single-layer can be found in e.g. Stefan Schaer [11].

Combined the Eqs. (21.5–21.7), the final observation equation of regional VTEC model is written as:

$$\sum_{i=0}^{n} \sum_{k=0}^{m} E_{ik} (\varphi - \varphi_0)^i (S - S_0)^k + 9.52437 \cos z' (dq_j + dq^i) = 9.52437 \cos z' (P^i_{2j} - P^i_{1j}) \tag{21.8}$$

Considering the much higher noise level of the ionosphere measurements derived from the code measurements, the carrier phase observations from L1 and L2 can be used to smooth the code observation for a more precise TEC estimate. Such carrier phase smoothing technique is also often referred as "carrier phase leveling", as mentioned in Hu et al. [12]. The carrier phase observation of the satellite are choose to be observed continuously when used to smooth the code observation.

In Eq. (21.8), the DCB of receiver and satellite are set as two separate parameters. Thus the unknown parameters can't be calculated in the principal of

least square because the normal equation is rank deficient. So the following limited equation is accepted (the same as CODE accepts):

$$\sum_{i=0}^{m} q^i = 0$$

where
m is all the number of satellite;
q is the DCB of the satellite;

In this paper, the orders of the three-dimensional Taylor series expansion in Eq. (21.8) are set to be 2 and 3. Suppose that one receiver observes k satellites during one period and there are t receivers. So the unknown parameters are 6 + k + t (6 coefficients, k satellite DCB, t receiver DCB). Usually the observations of one station in 2 h are more than 400, so the parameters can be calculated successfully in the principal of least square.

The data from JSCORS are used to evaluate the DCB of receiver and satellite. In this paper, the data is from the station of BTJH, from 10th March to 19th March, 2012. And the interval is 15s while the cut-off angle is 15°. Using the method mentioned above, the calculated DCB of the satellite are compared with the DCB that CODE issued in the following (the result of the early five days is shown, Fig. 21.1)

From Fig. 21.1, we find that most of the difference between the satellites' DCB calculated and issued are within 1 ns expect the satellite of PRN10, PRN22 and PRN31, which differ about 2 ns. The reason to cause the difference on one hand may be that the global harmony model which CODE accepts are different from regional polynomial model in this paper; on the other hand, CODE uses global data from all over the world whose observe time is much sufficient. Results show the standard deviation of the calculated DCB of satellite during the ten days is

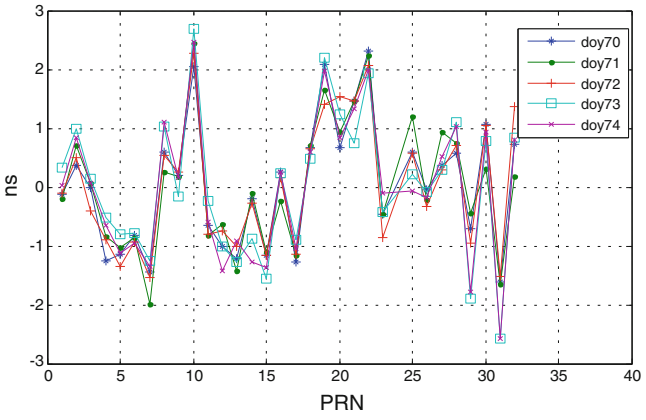


Fig. 21.1 Difference between DCB calculated and CODE issued

about 0.4 ns while receiver DCB is about 0.14 ns. The results are agreed to the conclusions made in Chang et al. [2] and Li et al. [13] that the DCBs are quite stable during one month time or more.

21.4 Real-Time TEC Model and Accuracy Analyses

According to the characteristic that the DCB of receiver and satellite are stable in a period, the real-time TEC model is established not only by taking DCB of the former day as known parameters into today's model, but also by ignoring the time dimension in Eq. (21.6), because the real-time model is established by single epoch.

The data in 6th February 2009 from three stations, BTRD, BTXH, BTYX of JSCORS is used to calculate the DCB of receiver and satellite. Then the DCB are taking into the real-time model of the next day (7th February) as known parameter to calibrate TEC. The interval of the data is 15s, and the cut-off angle is 15°. The real-time model is established every one minute. The internal RMS is calculated by the Eq. (21.10) in the following according to the difference between the observe value calculated by Eq. (21.5) that has map into the vertical direction (as VTEC) and the model value calculated by Eq. (21.6). Results show in Fig. 21.2.

$$RMS = \sqrt{\frac{V^T PV}{n-1}} \tag{21.10}$$

In order to check the accuracy of this method further, the data from the station of NJPK in the same epoch of the same day is used to test the external RMS information. The relative position of the stations is shown in Fig. 21.3. Only the ionosphere pierce point which is within the filed of the established model is counted while calculating the external RMS. The result shows in Fig. 21.4.

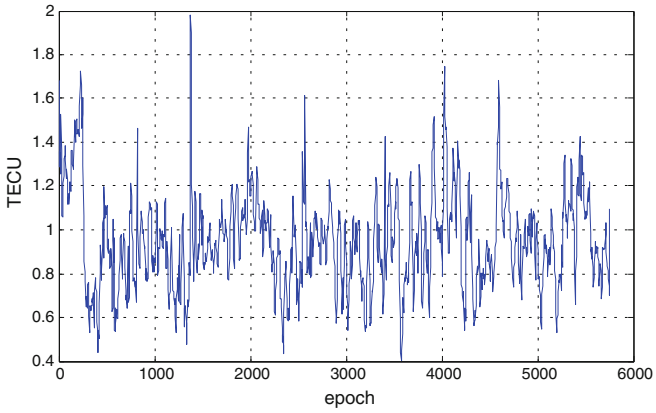


Fig. 21.2 Internal RMS accuracy

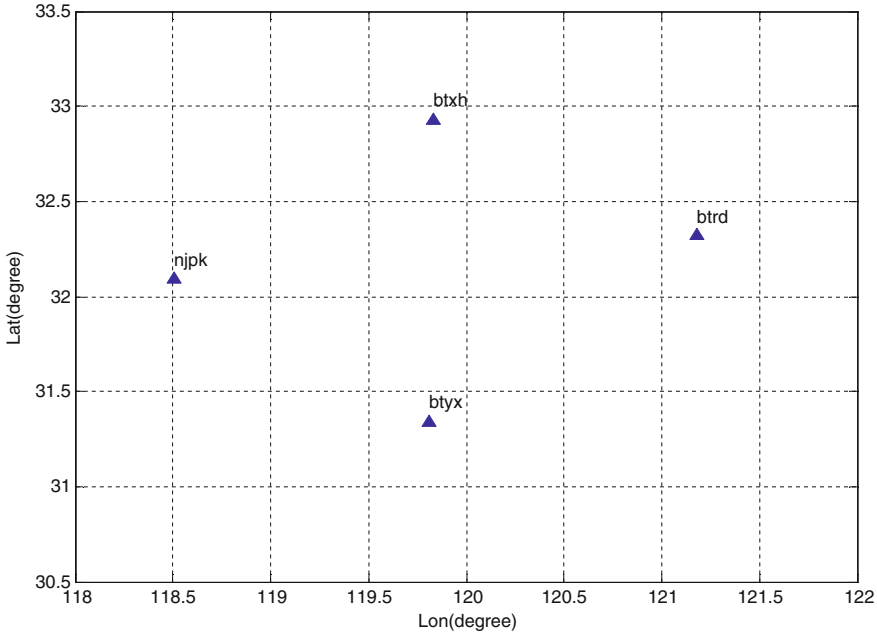


Fig. 21.3 Relative position of stations

The result indicates that the average internal and external RMS of the established real-time model with multi-station is 0.946 TECU and 0.948 TECU in the condition that the distance between the stations is about 150 km. The internal and external RMS in Figs. 21.2 and 21.4 show that most of them are about 1 TECU, but several reach 2–3 TECU. The reason may be that in those epoches, the latitude of some ionosphere pierce points is so small that deviate the center of the field of

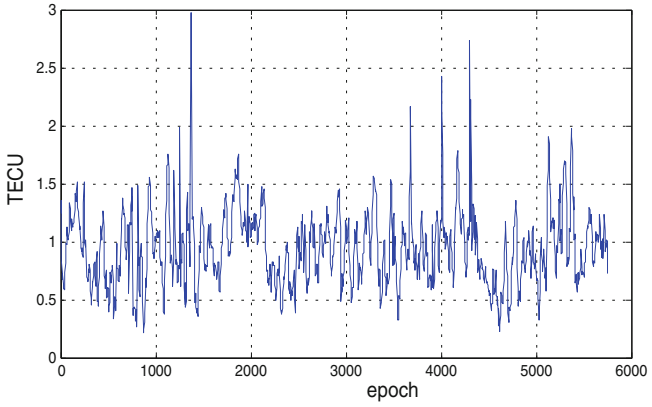


Fig. 21.4 External RMS precision

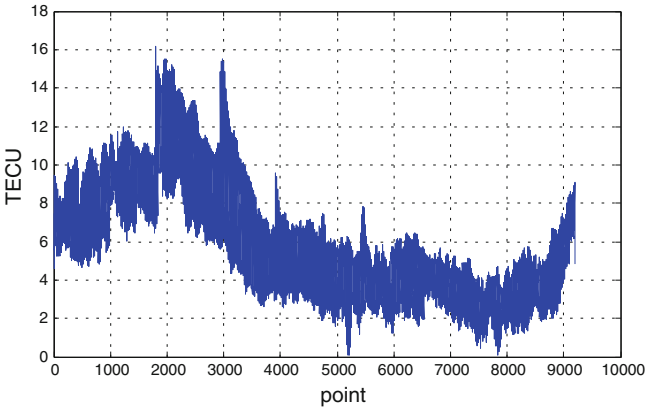


Fig. 21.5 Observe value

the established model. In those cases, the ionosphere pierce points are not so well-distributed that the accuracy of the model decreases. In conclusion, multi-station have advantages over single station in the real-time monitor of regional ionosphere, because multi-station possesses more ionosphere pierce points and thus makes the established area well-distributed. So the internal and external RMS accuracy improve much.

Figures 21.5 and 21.6 give the results of observe values and model values of all the ionosphere pierce points in all epoches during the whole day on 7th February 2009 of the checking station of NJPK.

From Figs. 21.5 and 21.6, we find that the difference between the model value calculated by the real-time model and observe value is little. The difference between the model value and observe value agrees to the result of the external RMS analyzed above. So diurnal variation of the TEC in the zenith direction of the station can be displayed by the real-time model. Figure 21.7 shows the 24-h change of ionospheric TEC in the zenith direction of NJPK station (33.053° 118.302°) on the day of 7th February 2009.

Convert UT time in Fig. 21.7 into local time, we can find that from about 6:30 a.m., TEC rises at the rise of the sun, and it reaches peak value at about 12:00–14:00 when the sun actives most. And then the amount of TEC decreases until the sun drop down. During the nighttime, TEC becomes quite steady.

So, to a certain area, we can monitor the 24-h change of TEC in real time by building corresponding grid net as the zenith direction of the grid point can be monitored. Moreover, the spatial and temporal characteristic of the regional ionosphere can be analyzed by the time series of TEC through those grid points around the area.

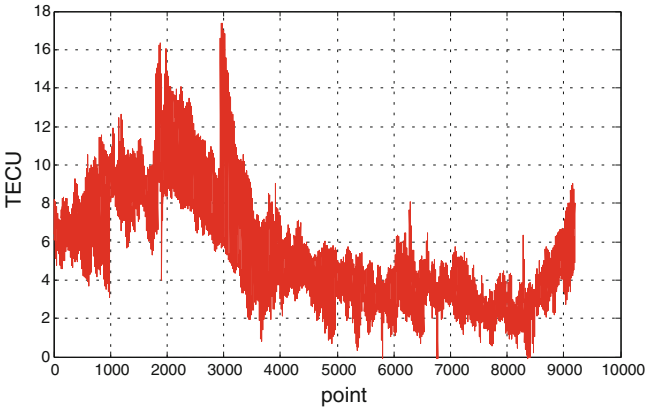


Fig. 21.6 Model value

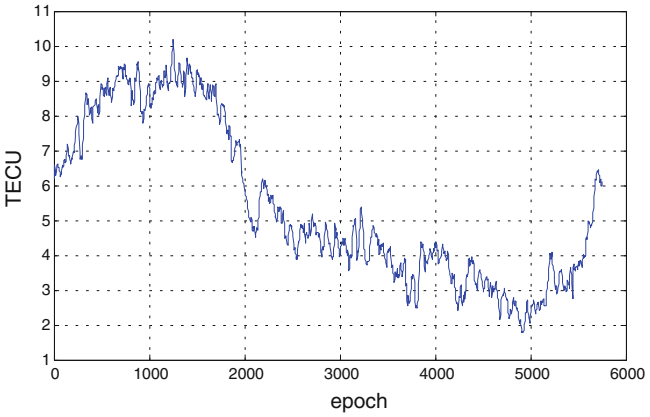


Fig. 21.7 TEC diurnal variation in the zenith direction of NJPK station

21.5 Conclusions

In this paper, the method of real-time monitor of ionospheric TEC based on polynomial model with multi-station is studied. Results show that the internal RMS of the model can be better than 1 TECU while the external RMS is about 1 TECU when the distance within the stations is about 150 km. The accuracy is much improved than the model with single station. Meanwhile, the TEC result of monitoring the zenith direction of certain station shows that the calculated time series of TEC can reflect the TEC change at the changing of the sun's activity. So this method can be used to monitor the TEC change of certain area in real time by building corresponding grid net.

Acknowledgments This study is funded and supported by National Natural Science Foundation of China (41074021, 41274028).

References

1. Yuan YB (2002) Study on theories and methods of correcting ionospheric delay and monitoring ionosphere based on GPS. Ph.D. thesis, Institute of Geodesy and Geophysics, Chinese Academy of sciences, Wuhan, China
2. Chang Q, Zhang DH, Xiao Z, Zhang QS (2001) A method for estimating GPS instrumental biases and its application in TEC calculation. Chin J Geophys, 44(5):596–601
3. Li GZ, Ning BQ, Yan H (2005) Design and realization of a real-time GPS ionospheric scintillation monitoring system. Chin J Radio Sci 20(6):758–764
4. Chen FH, Yuan H (2007) Study of method to compute ionospheric S4 index based on GPS signal. Chin J Space Sci 27(1):23–28
5. Lambert W (2004) Ionospheric disturbance indices for RTK and Network RTK positioning. In: Proceedings ION GNSS 2004, Long Beach, CA, 2849–2854
6. Aarons J, Mendillo M, Yantosca M (1996) GPS phase fluctuations in the equatorial region during the MISETA 1994 Campaign, J Geophys Res, 101(A12):26851–26862
7. Pi X, Mannucci AJ, Lindqwister UJ, Ho CM (1997) Monitoring of global ionospheric irregularities using the worldwide GPS network. Geophys Res Lett 24:2283–2286
8. Geng CJ, Tang WM, Zhang HP (2008) Real-time monitoring of ionosphere changes by CORS. J Geodesy Geodyn 28(5):105–108
9. Liu JN, Chen JY, Ge MR (1999) Principle and method of wide area differential GPS. Surveying and mapping press, Beijing
10. Zhang HP, Shi C, Tang WM (2008) United solution to polynomial VTEC modeling and DCB analysis using ground-based GPS observations. Geomatics Inf Sci Wuhan Univ 33(8):805–809
11. Stefan S (1999) Mapping and predicting the Earth's ionosphere using the global positioning system, Ph.D. thesis, Astronomical Institute, University of Berne, Switzerland
12. Hu WS, Gao CF (2002) Measuring principle of GPS and its application. Chin Commun Press, Beijiing
13. Li D, Wang J, Song SL, Lu XS (2011) Calculating the differential code bias based on the areal ionosphere model. J Surv mapping 20(6):40–44

Chapter 22
Analysis of the Reliability of the Protected Memories Affected by Soft Errors

Xiao-hui Liu, Hong-lei Lin and Wei-hua Mou

Abstract The formulations of mean time to failure and reliability under multiple factors are deduced of the soft errors of space memory systems affected by SEU (single event upsets). A quickly-to-calculate approximate formulas for mean time to failure of memories is produced. The optimal cost function which satisfied the system reliability is obtained by analyzed of the effect of the mean time to failure from four aspects: scrubbing interval, memory size, code word length and error correction capability of ECC (error correct codes). The results show that the memory reliability will be improved by reduced the scrubbing interval and increased the error correction capability of the ECC. The memory reliability will be changed by used of the ECC of different code word length, even though the ECC has the same correction capability. All above of the elements of the cost function will be considered to reach the system reliability. The results of this paper can be used to design other space memory systems.

Keywords Mean time to failure (MTTF) · Reliability · Error detection and correction (EDAC) · Single event upsets (SEUs)

22.1 Introduction

The soft errors of memories caused by Single event upsets (SEUs) are the main factor influencing the reliability of spacecraft [1]. In order to protect digital data against errors that can occur in memories, besides strengthening the hard wares (such as technologies or logical circuit, etc.), error detection and correction (EDAC) is usually used of system-level structure [2], i.e. detecting and correcting

X. Liu (✉) · H. Lin · W. Mou
Satellite Navigation R&D Center, School of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan, China
e-mail: lululiu_nudt@sina.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_22,

soft errors by error correct code (ECC), such as Hamming Code [3] for Single-error-correction double-error-detection (SEC-DED), and RS code [4] and Golay code [5] for correcting multiple-bit errors. EDAC is reading-decoding-correcting (scrubbing) in memories in set time by adding redundancy check of code. When a memory bit is hit by a high-energy event, 0/1 soft error appears. ECC can automatically detect and correct the error and cover the error with the correct data. However, EDAC needs extra CPU resource, which is relevant to ECC type, scrubbing interval and the size of the protected memories. Redundancy length added by ECC occupies the limited memories, and the redundancy length directly influences the reliability of memories. The above factors interact and influence the reliability of EDAC technique and the reliability of the whole memories.

When memories are protected by EDAC, the factors relevant to the system-level reliability are ECC type, scrubbing interval, protected memories size and space environment. Designers hope to reach high standard of reliability in low cost, i.e. longer scrubbing interval, simpler and better ECC and smaller protected memories. The reliability of memories by using EDAC has been studied by many researches [6–10]: Maestro and other people studied the reliability of memories under single-error correction (SEC) codes and deduced a statement that mean time to failure (MTTF) of memories with or without scrubbing. But the influence on reliability by multiple-bit errors correction codes was not shown in the statement. Meanwhile, Maestro also analyzed how disposition of memories influence reliability when using ECC and deduced a formula. But Maestro still failed to consider the contribution of scrubbing on reliability. Saleh and other people deduced a formula and emulated analysis on how three cases, i.e. no-scrubbing, deterministic scrubbing and probabilistic scrubbing, influence reliability of memories [8], however, their analysis is only about use of SEC codes.

On the basis of above researches, this paper deduces a formula on reliability and MTTF under multiple factors in EDAC. Quantitative analyzes are carried on related parameters in EDAC including error correction capability, scrubbing interval, memory size and code word length, etc. in order to clarify their influence on reliability independently or together. The paper also analyzes the influence of cost factor in the project and proposes a scheme of the optimal cost function satisfying reliability requirement, a reference to solution on optimal EDAC under reliability requirement in the project scheme.

22.2 MTTF Under Multiple Factors

The analysis of this paper is under the following assumptions: (1) soft errors of the memories occur with a Poisson distribution; (2) the errors in memories are statistically independent and linearly uncorrelated, i.e. close multiple simultaneous errors made by single event are not under the discussion; (3) the reliability of control, correction and interface circuitry in the memory is not under the discussion, i.e. the circuitry are fault-tolerant.

According the above assumptions, the study of memories reliability is calculating MTTF. For convenience, the variables are defined as follows:

M number of code words in memories
N total memory size
w number of data bits per code word
c number of check bits per code word
L error correction capability, i.e. ECC can correct L errors in a code word
λ event arrival rate per code word
λ_{bit} event arrival rate per bit, in this paper typical 1×10^{-6} upset/bit $\times$ day is used
T_S scrubbing interval

Define Wi is the fault time of the ith code word, and event arrival rate per code word λ follows Poisson distribution, then Wi is follows Gamma distribution of $\Gamma(L + 1, \lambda)$ [11], then probability density function is:

$$P_f(W_i = t) = \frac{(\lambda \cdot t)^L \mathrm{e}^{-\lambda \cdot t} \lambda}{L!} \tag{22.1}$$

The relationship between event arrival rate per code word and event arrival rate per bit is:

$$\lambda = \lambda_{bit}(w + c) \tag{22.2}$$

Then the fault probability of the ith code word in time t_x is:

$$P_f(W_i < t_x) = \int_0^{t_x} \frac{(\lambda \cdot t)^L \mathrm{e}^{-\lambda \cdot t} \lambda}{L!} dt \tag{22.3}$$

So, $\int x^n \mathrm{e}^{-ax} dx = -\frac{\mathrm{e}^{-ax}}{a^{n+1}}\left[(ax)^n + n(ax)^{n-1} + n(n-1)(ax)^{n-2} + \cdots + n!\right]$

Then integral calculus of (22.3) is:

$$\begin{aligned} P_f(W_i < t_x) &= \int_0^{\lambda t_x} \frac{x^L \mathrm{e}^{-x}}{L!} dx \\ &= -\frac{\mathrm{e}^{-\lambda \cdot t_x}}{L!}\left[(\lambda t_x)^L + L(\lambda t_x)^{L-1} + L(L-1)(\lambda t_x)^{L-2} + \cdots + L!\right] + 1 \end{aligned} \tag{22.4}$$

Then reliability of the ith code word in time t_x is:

$$r_i(t_x) = r(W_i < t_x) = 1 - P(W_i < t_x) = \frac{\mathrm{e}^{-\lambda \cdot t_x}}{L!}\left[(\lambda t_x)^L + L(\lambda t_x)^{L-1} + \cdots + L!\right] \tag{22.5}$$

If scrubbing interval is T_s, then the reliability of the ith code word in a scrubbing interval T_s is:

$$r_i(T_s) = \frac{e^{-\lambda \cdot T_s}}{L!}\left[(\lambda T_s)^L + L(\lambda T_s)^{L-1} + \cdots + L!\right] \tag{22.6}$$

If the memory consists of M code words, then the reliability of the whole memory in a scrubbing interval T_s is:

$$R(T_s) = r_i(T_s)^M = \frac{e^{-\lambda \cdot T_s M}}{L!^M}[(\lambda T_s)^L + L(\lambda T_s)^{L-1} + \cdots + L!]^M \tag{22.7}$$

So MTTF of memory T_{MTTF} is:

$$T_s\left[1 + R(T_s) + R^2(T_s) + \cdots\right] < T_{MTTF} < T_s\left[R(T_s) + R^2(T_s) + R^3(T_s) + \cdots\right]$$

Usually $T_{MTTF} \gg T_s$, the approximate expression is:

$$T_{MTTF} \approx \frac{T_s}{1 - R(T_s)} \tag{22.8}$$

Put (22.7) into the above expression:

$$T_{MTTF} = \frac{T_s}{1 - \frac{e^{-\lambda \cdot T_s M}}{L!^M}\left[(\lambda T_s)^L + L(\lambda T_s)^{L-1} + \cdots + L!\right]^M} \tag{22.9}$$

22.3 Influences of Parameters on MTTF

22.3.1 The Influence of Code Word Length on MTTF

If the error correction capability L is fixed, the data length of a code word is increased k times and check length is c_k, there will be not a relationship of k times between c_k and c. When the new ECC is used (data length increased k times, error correction capability is unchanged), ratio of MTTF before and after change is ξ_{kw}. From the definition of arrival ratio of single event in a code word, we know $\lambda = \lambda_{bit}(w + c)$. Once data length is increased k times, arrival ratio of single event in a code word is $\lambda_k = \lambda_{bit}(kw + c_k)$. Define the expression $\varepsilon(\lambda) = \frac{(\lambda T_s)^L}{L!} + \cdots + \frac{L(L-1)\cdots 2}{L!}(\lambda T_s)$ to be the function of λ, then

$$\xi_{kw} \approx \frac{M_k \lambda_k T_s[1 + \varepsilon(\lambda_k) - \varepsilon(\lambda_k)/(\lambda_k T_s)]}{M\lambda T_s[1 + \varepsilon(\lambda) - \varepsilon(\lambda)/(\lambda T_s)]} \tag{22.10}$$

Two situations will be discussed as follows:

1. If the total protected data size ($w{\cdot}M$) is unchanged, after using the new ECC, number of code word is $M_k = M/k$. $\eta = w/(w + c)$ and $\eta_k = kw/(kw + c_k)$ are code efficiency before and after the change of ECC, then (22.10) can be simplified as:

$$\xi_{kw} = k(\eta/\eta_k)^2 \tag{22.11}$$

   The above expression shows that when the protected data size is unchanged and the same error correction capability, if the data length is changed, the MTTF is relevant to k and the ratio of code efficiency.
2. When total memory size N is unchanged, data length is w and kw, code word number in memory is $M = N/(w + c')$ and $M_k = N/\left(kw + c'_k\right)$. The length of c' and c'_k is the smallest round number times bigger than c and c_k. Calculate (22.10):

$$\xi_{kw} = \frac{(1 + c'/w)(k + c_k/w)^2}{(k + c'_k/w)(1 + c/w)^2} \tag{22.12}$$

The data length of ECC increased to k times, its MTTF ratio of before and after the change ξ_{kw} is a function having nothing to do with L, i.e. having nothing to do with the error correction capability of the used ECC. When $k = 1$ (i.e. data length unchanged), higher code efficiency (redundancy length shorter) is, longer MTTF. When data length increases ($k > 1$), longer the code length is, shorter MTTF. However, when L is fixed, ECC whose code length is longer, its code efficiency is higher, i.e. when memory size is fixed, the ratio that effective data length occupies redundancy check length. So there should be a compromise between data size and MTTF when codes are designed. The below table shows that in order to use ECC (13, 8), when error correction capacity $L = 1$ and $L = 2$, when ECC is changed, after using code (22, 16), (39, 32) and (72, 64), the ratio of MTTF before and after the change. The result of calculation expresses that using ECC whose code length is longer, although code efficiency is improved and memory size to put effective data length is improved, as L is unchanged, as code length increases and probability that code is influenced by single event increases, the MTTF of whole memory is decreased.

22.3.2 The Influence of Scrubbing Interval on MTTF

Usually scrubbing can improve the reliability of memory, so scrubbing interval is a parameter importantly designed in the project. Generally speaking when scrubbing interval is infinitely small, reliability of memory can reach 1. In fact this is not the truth. Highly frequent scrubbing might waste lots of hardware resource and also

hard to be realized. The sensible choice should be suitable scrubbing interval which can satisfy reliability standard of system on the condition of hardware.

If other parameters unchanged, scrubbing interval is α ($\alpha > 0$) times (i.e. αT_s), the ratio of MTTF before and after the change is $\xi_{\alpha T_s}$. If the function T_s is $\beta(T_s) = \frac{(\lambda T_s)^L}{L!} + \frac{L(\lambda T_s)^{L-1}}{L!} + \cdots + (\lambda T_s)$, then

$$\xi_{\alpha T_s} = \frac{1 - e^{-\lambda\alpha T_s M}[1 + \beta(\alpha T_s)]^M}{\alpha\{1 - e^{-\lambda T_s M}[1 + \beta(T_s)]^M\}} = \frac{1 - [(1 - \lambda\alpha T_s)(1 + \beta(\alpha T_s)]^M}{\alpha\{1 - [(1 - \lambda T_s)(1 + \beta(T_s)]^M\}} \quad (22.13)$$

Because $|\lambda T_s| \ll 1$, and usually the change of scrubbing interval α is a finite number, then $|\lambda\alpha T_s| \ll 1$, and $|\beta(T_s)| \ll 1$, $|\beta(\alpha T_s)| \ll 1$ and $M > 0$. Developing the above formula in Taylor method and calculate it, we can get an approximate expression:

$$\xi_{\alpha T_s} \approx \frac{(\lambda\alpha T_s)^2}{\alpha(\lambda T_s)^2} = \alpha \quad (22.14)$$

According to above expression, $\xi_{\alpha T_s}$ is a function having nothing to do with error correction capacity L and scrubbing interval T_s. When other parameters fixed, change T_s into α times, MTTF will be $1/\alpha$ times. When scrubbing interval is decreased, i.e. $\alpha < 1$, MTTF will increase to $1/\alpha$ times. When scrubbing interval is increased, i.e. $\alpha > 1$, MTTF will decrease to $1/\alpha$.

22.3.3 The Influence of Error Correction Capacity on MTTF

When error correction capacity changes and data length is unchanged, the related check length of ECC will change, if the whole memory size N is unchanged, code word number M will be changed. If error correction capacity is improved from L to $L + 1$, code word number changes from M_L to M_{L+1}, ignoring the change of event arrival rate λ, the ratio of MTTF after and before the change is function $\xi_L.\gamma(L) = \frac{(\lambda T_s)^L}{L!} + \frac{L(\lambda T_s)^{L-1}}{L!} + \cdots + \lambda T_s$ is the function of error correction capacity L. Then

$$\xi_L = \frac{[1 + \gamma(L) - \gamma(L)/\lambda T_s - \lambda T_s/2 - \lambda T_s\gamma(L)/2]M_L}{[1 + \gamma(L + 1) - \gamma(L + 1)/\lambda T_s - \lambda T_s/2 - \lambda T_s\gamma(L + 1)/2]M_{L+1}} \quad (22.15)$$

Figure 22.1 shows when error correction capacity changes from 1 to 2 ($Ratio_1$), from 2 to 3 ($Ratio_2$), and from 3 to 4 ($Ratio_3$), how MTTF changes. The figure shows as error correction capacity improves, MTTF also improves. But as scrubbing interval increases, reliability brought by strong error correction capacity decreases.

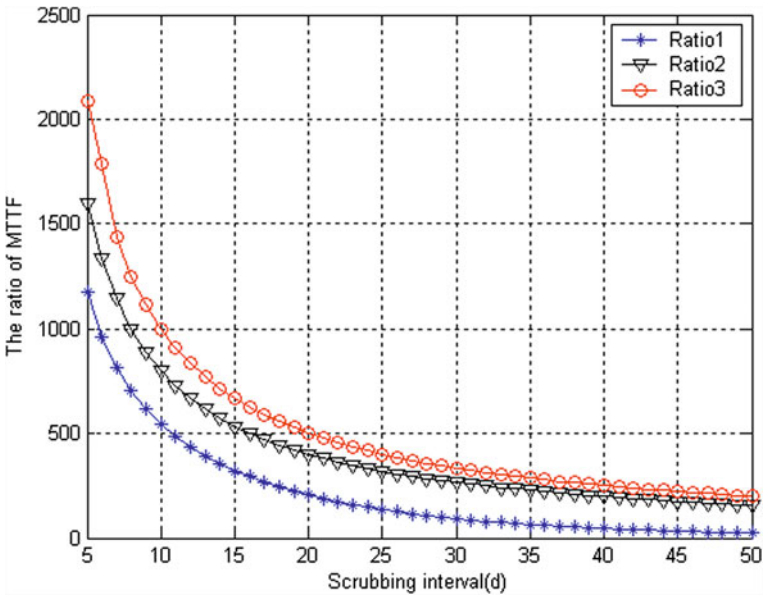


Fig. 22.1 The relationship of error correct capability and the ratio of MTTF

22.4 The Design and Analysis of Cost Function

Basing on the above analysis of influences on reliability from several factors, we hope to find a scheme with long scrubbing interval, easy code design and high system reliability. From above analysis, aims can not be realized together. Obviously, this is a problem of Multi-objective Optimization Problem (MOP). V.Pareto who studied MOP in economics, proposed a concept of Pareto [12, 13].

22.4.1 Description of Problem

22.4.1.1 Design of Variables

The optimal design in this paper is on the condition of eight different ECC that error correction capacity is 1 and 2 (see Table 22.1), as $X = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8]$, so variables can be designed as $x_i = (w_i,\ c_i)$, w_i is data length, c_i is check length. Based on influences on reliability from several factors, we design cost function to realize EDAC in order to minimize cost while satisfying restricted conditions.

Table 22.1 Ratio of MTTF for ECC of different code word length

	L = 1		L = 2	
	MTTF	Check bits	MTTF	Check bits
8bit	1	5	1	9
16bit	1.91	6	1.89	11
32bit	3.60	7	3.50	13
64bit	6.81	8	6.48	15

22.4.1.2 Description of Optimal Aims

Complexity of Decode

Usually delay time is used to measure complexity of decode, i.e. time of decoding in different ECC to measure complexity $f(w)$. We always hope to choose less complex program to satisfy anticipated standards, i.e. aim can be designed as $f_1(x) = f(w)$.

Utilization Ratio of Memory Size

In the same data sizes, using the code with long data length, it will decrease redundancy memory size. Utilization ratio of memory size in different code efficiency $g(x)$ can be used as an index, we always hope to use resources at most when designing scheme. So the scheme whose utilization ratio of memory size is higher is chosen, i.e. aim can be designed as $f_2(x) = 1/g(x)$.

Scrubbing Interval

The factor of scrubbing interval consists of time to scrub t_s and scrubbing interval T_s. In actual system $T_s \geq t_s$, t_s is direct ratio of complexity of code and memory size. The shorter scrubbing interval, frequency of memory scrubbing is higher. Obviously in this situation the cost of the system is larger. So we always hope to choose the program whose scrubbing interval is bigger, then aim can be designed as $f_3(x) = 1/T_s(x)$.

22.4.1.3 Restricted Conditions

Restricted conditions are based on the requirement of design. The optimal choice of ECC in this paper is based on the requirement of reliability and memory size. So main restricted conditions are reliability and memory size, and memory size can be divided into total memory size and protected information size. The paper focuses on the second situation, the first situation can be drawn a similar conclusion.

When total protected information size is limited, the restricted condition is $T_{MTTF_i} \geq T_{MTTF_0}$, and $M_i w_i \leq N_0'$. Here T_{MTTF_0} is required MTTF in system, N_0' is protected information size.

22.4.1.4 Multi-Objective Optimization Model

Analyzing optimal aims and restricted conditions, the design program of EDAC in memory can be:

$$\begin{cases} find & x \in X \\ \min & (f_1(x), f_2(x), f_3(x, T_{MTTF}, M, L))^T \\ st. & T_{MTTF_i} \geq T_{MTTF_0} \\ & M_i w_i = N_0' \; i = 1, \ldots 8 \end{cases} \tag{22.16}$$

22.4.2 Solution to the Problem

22.4.2.1 Way of Solving

In order to solve MOP, the paper adopts a usual method, i.e. transforming solution to MOP into solution to single objective optimization problem with a new comprehensive aim function made by linear weighted averages method. ω_1 ω_2 ω_3 are complexity of code, memory size and scrubbing interval, weighted factors in comprehensive function $F(x)$, the expression is $F(x) = \sum_{j=1}^{3} \omega_j f_j(x)$. So the optimal problem in (22.16) can be transformed as

$$\begin{cases} find & x \in X \\ \min & F(x) = \sum_{j=1}^{3} \omega_j f_j(x) \\ st. & T_{MTTF_i} \geq T_{MTTF_0} \\ & M_i w_i = N_0' i = 1, \ldots 8 \\ & \sum_{\mathrm{j}=1}^{3} \omega_{\mathrm{j}} = 1 \omega_j \geq 0 \end{cases} \tag{22.17}$$

22.4.2.2 Solution to Aims

Aim of Complexity of Decode $f_1(x)$

$f_1(x_i)$ is delay time of the x_{i} type ECC of a code being read-decoded-written. Exact time figures in emulation experiments by anti-fuse FPGA AX1000-2-F896 are shown in Table 22.2.

Aim of Utilization Ratio of Memory Size $f_2(x)$

Utilization ratio of memory size is the ratio of information length to total memory size, after normalizing utilization ratio of memory size $g(x)$ is shown in Table 22.3.

Table 22.2 Delay time and the normalize factor

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
ECC	(13, 8)	(22, 16)	(39, 32)	(72, 64)	(17, 8)	(27, 16)	(45, 32)	(79, 64)
Delay time (ns)	32	45	60	78	38	51	63	83
Normalize factor	0.39	0.54	0.72	0.94	0.46	0.61	0.76	1

Table 22.3 Normalize the capability of the memory

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$g(x)$	0.56	0.75	0.9	1	0.38	0.56	0.75	0.9

Table 22.4 Scrubbing interval of the different ECC under the same MTTF

L = 1	Scrubbing interval (s)	L = 2	Scrubbing interval (s)
(13, 8)	4.735	(17, 8)	703.4375
(22, 16)	2.4798	(27, 16)	354.2317
(39, 32)	1.3151	(45, 32)	200.9246
(72, 64)	0.6945	(79, 64)	111.293

From $f_2(x) = 1/g(x)$, the function of utilization ratio of memory size $f_2(x)$ can be calculated.

Scrubbing Interval Aim in Same Restricted Conditions $f_3(x)$

In order to get optimal result, we should calculate scrubbing intervals in different ECC satisfying restricted conditions T_{MTTF_0} and N_0'. Table 22.4 shows scrubbing intervals in different ECC when $T_{MTTF_0} = 10^5$ (d) $N_0' = 1$ MByte.

22.4.3 Optimal Result and Its Analysis

22.4.3.1 Influence of Weighted Vectors on Optimal Results

When we transform the solution to MOP into the solution to single objective optimization problem by linear weighted averages method, weighted factors are greatly influenced by objective factors. In order to describe the influence of weighted vectors on optimal result objectively, the paper emulates optimal results in different weighted vectors when restricted conditions are $T_{MTTF_0} = 10^4(d)$ and $N_0^{'} = 1$ MByte. The relationship between optimal solutions and weighted vectors is shown in Fig. 22.2.

According to Fig. 22.2, when weighted coefficient scrubbing interval is bigger, choosing the ECC with strong error correction capacity can decrease cost. When weighted coefficient utilization ratio of memory size is bigger, choosing the ECC with long code length can decrease cost of the scheme. When weighted coefficient

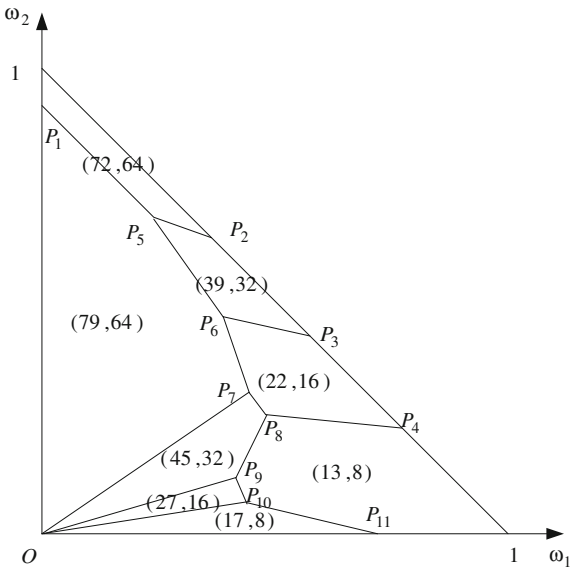


Fig. 22.2 The relationship of weight vectors and optimal results area

code complexity is bigger, choosing the ECC with short code length and weak error correction capacity can decrease cost of the program. The values of weighted factors (ω_1, ω_2) in the boundary are shown in Table 22.5.

According to Table 22.5, when weighted vectors are known, we can gain optimal design directly and design the values of weighted factors.

22.4.3.2 Influence of Reliability Restricted Condition on Optimal Result

When analyzing the cost changing with the reliability restricted condition, we can fix a group of weighted coefficient, i.e. when weighted vectors $(\omega_1, \omega_2, \omega_3)$ are known, we discuss the changing relationship between comprehensive aim function $F(x)$ in the model and T_{MTTF}. With emulation, we choose weighted factor vectors (0.5, 0.4, 0.1), T_{MTTF} change from 103 to 105, the result is shown in Fig. 22.3.

Table 22.5 Weight factors of the optimal result's area boundary

	(ω_1, ω_2)		(ω_1, ω_2)		(ω_1, ω_2)
P1	(0, 0.96)	P2	(0.33, 0.67)	P3	(0.55, 0.45)
P4	(0.74, 0.26)	P5	(0.21, 0.75)	P6	(0.42, 0.52)
P7	(0.44, 0.48)	P8	(0.60, 0.29)	P9	(0.62, 0.20)
P10	(0.69, 0.12)	P11	(0.92, 0)		

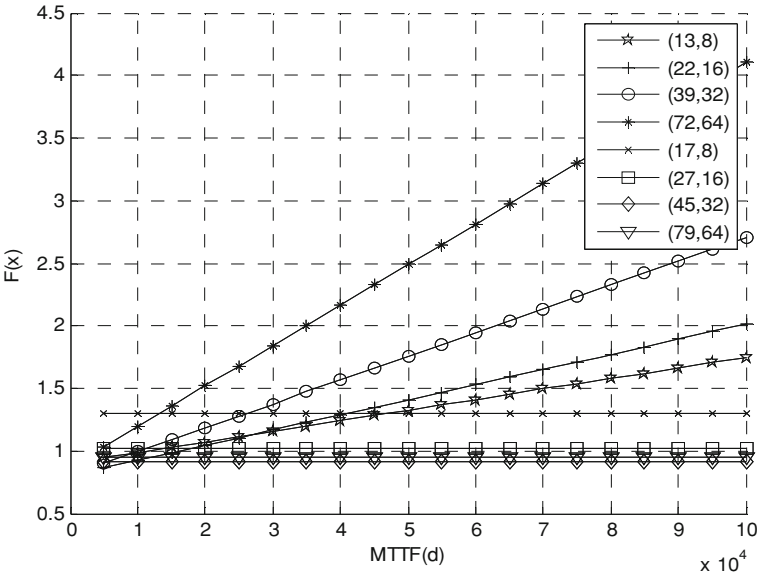


Fig. 22.3 The relationship of generalized objective function and MTTF

From the emulation result, when $T_{MTTF} < 1.5 \times 10^4$, i.e. reliability requirement is low, choosing the ECC with weak error correction capacity can decrease the cost; on the contrary, choosing the ECC with strong error correction capacity when reliability requirement is high.

Comparing four ECC with error correction capacity $L = 1$ in emulation, with same error correction capacity, as increasing reliability requirement, choosing short code length cost less. Similarly, when $L = 2$, this conclusion can also be drawn.

22.5 Conclusions

According to above discussions on optimal designs of protected memory, we can draw the following conclusions.

1. Both decreasing scrubbing interval and increasing error correction capacity of ECC can improve the reliability of protecting memory; when scrubbing interval is changed from T_s to $1/\alpha$ ($\alpha > 1$) times, MTTF will be α times; with same code lengths, when error correction capacity improves, MTTF will be $f(T_s, M, L)$ times, f is non-linear function among scrubbing interval T_s, code number M and error correction capacity L, when T_s increases, f will decrease.
2. With unchanged error correction capacity, ECC type also influences reliability: when protected total data size is unchanged, using the code whose data length is

k times to the original, ratio of MTTF before to after the change $\xi_{kw} = k(\eta/\eta_k)^2$ is the function between k and code efficiency. If protected memory size is unchanged and data length of ECC is increased to k times, when $k = 1$ (i.e. data length unchanged), higher code efficiency (smaller redundancy length), longer MTTF; when data length increases ($k > 1$), longer code length, shorter MTTF.

3. Adopting EDAC to protect memory, we should consider influences of scrubbing interval, code error correction capacity, decode efficiency and other factors on system reliability and cost. Choosing different weighted factor leads to different optimal design: when decode complexity weighted factor is bigger, choosing short code length and the code with weak error correction capacity can decrease the cost; when memory size weighted factor is bigger, choosing long code length can decrease the cost; when scrubbing interval weighted factor is bigger, choosing the code with strong error correction capacity can decrease the cost.

References

1. Blake JB, Mandel R (1986) On-orbit observations of single event upsets in Harris HM-6508 IK RAMs [J]. IEEE Trans Nucl Sci 33(6):1616
2. Shirvani PP, Saxena NR, McCluskey E (2000) Software-implemented EDAC protection against SEUs. IEEE Trans Reliab 49(3):273–283
3. Zhang Yu, Yang-ming Zheng, Zheng-liang Huang et al (2008) Fault-tolerant design of memory module for pica-satellite on-board computer. J Astronaut 29(6):2057–2061
4. Bajura MA, Boulghassoul Y, Naseer R et al (2007) Models and algorithmic limits for an ECC-based approach to hardening sub-100 nm SRAMs. IEEE Trans Nucl Sci 54(4):935–945
5. Neuberger G, Lima F, Carro L et al (2003) A multiple bit upset tolerant SRAM memory. ACM transactions on design automatic electronic system
6. Maestro JA, Reviriego P (2009) Reliability of single-error correction protected memories. IEEE Trans Reliab 58(1):193–201
7. Maestro JA, Reviriego P (2009) Selection of the optimal memory configuration in a system affected by soft errors [J]. IEEE Trans Device Mater Reliab 9(3):403–411
8. Saleh AM, Serrano JJ, Patel JH (1990) Reliability of scrubbing recovery-techniques for memory systems. IEEE Trans Reliab 39(1):114–122
9. Zhu M, Xiao LY, Liu C et al (2011) Reliability of memories protected by multi-bit error correction codes against MBUs. IEEE Trans Nucl Sci 58(1):289–295
10. Reviriego P, Argyrides C, Maestro JA et al (2011) Improving memory reliability against soft errors using block parity. IEEE Trans Nucl Sci 58(3):981–986
11. Reviriego P, Holst L, Maestro JA (2010) Number of events and time to failure distributions for error correction protected memories. IEEE Trans Device Mater Reliab 10(3):381–389
12. Carls A (2000) An update survey of multi-objective optimization techniques. ACM Comput Surv 32(2):109–143
13. Marler RT, Arora JS (2004) Survey of muti-objective optimization methods for engineering. Struct Multi Optim 26(6):369–395
14. Goodman RM, Sayano M (1991) The reliability of semiconductor RAM memories with on-chip error-correction coding. IEEE Trans Inf Theor 37(3):884–896
15. Yang GC (1995) Reliability of semiconductor RAMs with soft-error scrubbing techniques. IEEE Proc Comput Digital Tech 142(5):337–344

Chapter 23
Research on Receiver Autonomous Integrity Monitoring Technology for Rocket Onboard GNSS Receiver

Rui Su, Aihua Xu, Shuqiang Zhao, Wei Zhang and Tongyu Zhang

Abstract The Receiver Autonomous Integrity Monitoring (RAIM) Technology can make the fault detection and identification with the redundant Global Navigation Satellite System (GNSS) information independently. In this paper, the RAIM technology is investigated for the high-speed rocket onboard GNSS receiver. And the RAIM algorithm for position domain is extended to the velocity domain. Moreover, the algorithm of outliers' elimination for pseudo range and Doppler frequency observation is proposed. The experiment results prove that the method can improve the accuracy and reliability of the rocket trajectory. Especially, it can also remove the velocity fluctuation of the rocket effectively.

Keywords RAIM · Rocket onboard GNSS receiver · GPS

23.1 Introduction

Precise rocket ballistic determination with GNSS onboard receiver is one of the significant tasks and approaches in the external trajectory measurement, which play an important role for the security of spacecraft experiments and improvement of the launch vehicle technology. Compared with the other GNSS positioning devices, the GNSS onboard receiver has some features such as higher speed and high dynamic range. Moreover, frequent unlock of the GNSS signal may result in the decline of the precision of position and velocity measurement.

With the redundant GNSS satellites information, the Receiver Autonomous Integrity Monitoring (RAIM) technology can make the Fault Detection (FD) and Fault Exclusion (FE) in positioning independently. Without any extra auxiliary

R. Su (✉) · A. Xu · S. Zhao · W. Zhang · T. Zhang
State Key Laboratory of Astronautic Dynamics, Xi'an Satellite Control Center, Xi'an, China
e-mail: ssurui@gmail.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_23,

equipment, RAIM has a lower cost and an easier way to achievement than the other methods, including the Wide Area Augmentation System (WAAS) and the Ground Augmentation Integrity Monitoring (GAIM). Hence, RAIM is a popular integrity monitoring method and a more appropriate choice for the rocket GNSS data processing.

Bi [1] investigates the correlation distance based RAIM theory and proposed a new algorithm in which the clustering analysis is used. Ying et al. [2] also introduce the parity space vectors (PSV) based RAIM theory and provide a simulation. Xu et al. [3] propose a method for calculating the threshold value of clock bias. According to the solved values of clock bias over a past period of time, a quadratic function model is established to get the value of predicted clock bias in current time. On the other hand, the estimated value of clock bias in current time is obtained from the measurements in current time. After the difference between the predicted value and the estimated value of clock bias in current time is got, a comparison is made between the difference and the threshold value of clock bias to accomplish the function of fault detection. Ren et al. [4] introduce two RAIM availability analysis methods including the maximum Horizontal Dilution of Precision (HDOP) and the Approximate Radial-error Protected (ARP). But these RAIM methods above are designed for positioning, Wang et al. [5] use the traditional RAIM algorithm for velocity domain. In their paper, the forecasting precision of clock bias is used in the integrity monitoring of velocity to satisfy the various RAIM requirements. However, the algorithm has a high computational complexity and mainly depends on the clock bias so that it not suitable to the rocket GNSS process.

In view of the features (high-speed and high dynamic range) of the rocket trajectory and the computational complexity, the RAIM technology is firstly investigated for the high-speed rocket onboard GNSS receiver in this paper. The correlation coefficient of the parity space vectors are used for the RAIM of the rocket positioning. Moreover, the PSV RAIM algorithm is extended to the velocity determination. By this method, the unreasonable pseudo range and Doppler frequency observations can be separated and eliminated according to the residual vectors of observations so as to improve the positioning accuracy and eliminate the velocity fluctuation effectively.

Finally, the algorithm is verified by the actual measurement data in the real launching tasks. The experiment results prove that the proposed RAIM methods can improve the accuracy and reliability of the rocket trajectory distinctly. Especially, it can remove the velocity fluctuation of the rocket effectively and make the precision of the rocket velocity much better.

23.2 PSV Based Fault Detection and Fault Exclusion

23.2.1 Single Point Position

With the theory of the pseudo range positioning, the positioning mode can be acquired [6], as below.

$$\mathbf{V} = \mathbf{B}\delta\mathbf{x} - \mathbf{L} \tag{23.1}$$

So the constant item $\mathbf{L}$ can be represented as flow.

$$\mathbf{L} = \mathbf{B}\delta\mathbf{x} - \mathbf{V} \tag{23.2}$$

If $n > 4$, the tiny correction $\delta\mathbf{x}$ of the receiver can be acquired with the Least Square Method (LMS) by the following equation.

$$\delta\mathbf{x} = (\mathbf{B}^{\mathbf{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathbf{T}}\mathbf{L} \tag{23.3}$$

Repeating the calculation of $\delta\mathbf{x}$ and correction of the skeleton coordinates of the receiver until $\delta\mathbf{x}$ is close to 0, the position of the receiver will be acquired.

23.2.2 Single Point Velocity Determination

After the derivation of the pseudo range positioning equation for time and the linearization, the equation can be written as the matrix form as blow.

$$\dot{\mathbf{V}} = \mathbf{B}\dot{\mathbf{X}} - \dot{\mathbf{L}} \tag{23.4}$$

It can be observed that the matrix $\mathbf{B}$ in equation (23.4) and the matrix $\mathbf{B}$ in equation (23.4) are same. The geometry relation of receiver and the GPS satellites is represented by the matrix $\mathbf{B}$. If $n > 4$, the velocity $\dot{\mathbf{X}}$ of the receiver can be acquired with LMS by the following equation.

$$\dot{\mathbf{X}} = (\mathbf{B}^{\mathbf{T}}\mathbf{B})^{-1}\mathbf{B}^{\mathbf{T}}\dot{\mathbf{L}} \tag{23.5}$$

23.2.3 RAIM for Rocket Positioning

From the theory of the single point position, it can be conclude that the position of the receiver is determined by the n GPS satellites. If the pseudo range observation of some GPS satellite contains a large bias, the position of the receiver calculated will have a large offset so that the positioning precision will be deceased.

Because the rocket usually flights in comparative high space and has a high speed, in addition with a high dynamic range, the GNSS rocket receiver has the following feature. One is that it can receive more signals of the GPS satellites than a common aground GNSS receiver. The other is that the pseudo range observation of some emerging GPS satellite usually contains a large bias. These features may cause that the positioning precision decease drastically in short time range. To overcome the flaw, the correlation coefficients of PSV are used for eliminating the unreasonable pseudo range observations in order to improve the positioning precision of the rocket.

The matrix $\mathbf{B}$ is processed by the Orthogonal-Triangular (QR) decomposition as the following equation.

$$\underset{n\times 4}{\mathbf{B}} = \underset{n\times n}{\mathbf{Q}} \quad \underset{n\times 4}{\mathbf{S}} \tag{23.6}$$

where the matrix $\mathbf{Q}$ is Orthogonal, that is $\mathbf{Q^T Q} = \mathbf{E}$. And the matrix $\mathbf{Q^T}$can be written as a partitioning form as blow.

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q_x} \\ \mathbf{Q_y} \end{pmatrix} \tag{23.7}$$

where the dimension of the matrix $\mathbf{Q_y}$ is $(n-4)\times n$. The QR parity transform matrix is defined as $\mathbf{T} = \mathbf{Q_y}$, and the QR parity vector $\mathbf{t}$ is also defined with the following equation.

$$\mathbf{t} = \mathbf{TL} \tag{23.8}$$

The constant item vector $\mathbf{L}$ can be substituted with the vector of $(\mathbf{B}\delta\mathbf{x} - \mathbf{V})$. Because the matrix $\mathbf{T}$ is orthogonal with the matrix $\mathbf{B}$ [1], the parity vector $\mathbf{t}$ change to the following form.

$$\mathbf{t} = \mathbf{TL} = \mathbf{T}(\mathbf{B}\delta\mathbf{x} - \mathbf{V}) = -\mathbf{TV} \tag{23.9}$$

The vector $\mathbf{V}$can be calculated as blow.

$$\mathbf{V} = (\mathbf{B}(\mathbf{B^T B})^{-1}\mathbf{B^T} - \mathbf{E})\mathbf{L} \tag{23.10}$$

Here define $\mathbf{T_i} = [\,T_{1i} \quad T_{2i} \quad \cdots \quad T_{mi}\,]^{\mathbf{T}}$, where $m = n - 4$. With $\varepsilon = -\mathbf{V}$, the expansion form of Eq. (23.10) can be as the following.

$$\mathbf{t} = \mathbf{T}\varepsilon = \varepsilon_1\mathbf{T_1} + \ \varepsilon_2\mathbf{T_2} \ + \ \cdots \ + \ \varepsilon_n\mathbf{T_n} \ (i = 1, 2, \cdots, n) \tag{23.11}$$

The vector $\mathbf{T_i}$ is determined by the geometry relation of the GPS satellites positions. And ε_i depends on the pseudo range observation. Hence, the vector $\varepsilon_i\mathbf{T_i}$ is jointly determined by the observation error and the satellites geometry. When the observation of satellite i contain a large bias, the vector $\varepsilon_i\mathbf{T_i}$ will usually be in the ascendant in $\mathbf{t}$, that is, the correlation coefficient of $\varepsilon_i\mathbf{T_i}$ and $\mathbf{t}$ will be large.

The matrix $\mathbf{A}$ is formed by merging together the vectors in Eq. (23.12), as blow.

$$\mathbf{A} = [\varepsilon_1 \mathbf{T_1} \;\; \varepsilon_2 \mathbf{T_2} \;\; \cdots \;\; \varepsilon_n \mathbf{T_n} \;\; \mathbf{t}] = [\mathbf{a_1} \; \mathbf{a_2} \; \cdots \; \mathbf{a_n} \; \mathbf{a_{n+1}}] \quad (i = 1, 2, \cdots, n) \tag{23.12}$$

The correlation coefficient c_i of the first n vectors and the last vector in the matrix $\mathbf{A}$ is defined as blow.

$$c_i = \frac{\sum_{k=1}^{m} (a_{k,i} - \bar{a}_i)(a_{k,n+1} - \bar{a}_{n+1})}{\sqrt{(\sum_{k=1}^{m} a_{k,i} - \bar{a}_i)^2 (\sum_{k=1}^{m} a_{k,n+1} - \bar{a}_{n+1})^2}} \quad (i = 1, 2, \cdots, n) \tag{23.13}$$

The c_i is more close to 1, it is more possible that the pseudo range observation of GPS satellite i contain a large bias. When couples of errors happen in the observations, the vector $\mathbf{t}$ will be influenced by these errors.

23.2.4 RAIM for Velocity Determination

The velocity of the receiver is also determined by the n GPS satellites. If the Doppler frequency observation of some GPS satellite contains the large bias, the velocity of the receiver calculated will have a large offset which usually cause the precision curve fluctuating in a period. To overcome the problem, the RAIM algorithm is extended to the velocity field.

According Eq. (23.6), the matrix $\mathbf{B}$ can be made the QR decomposition. So the matrix $\mathbf{T}$ is acquired also. QR parity vector $\dot{\mathbf{t}}$ for velocity can be defined as blow.

$$\dot{\mathbf{t}} = \mathbf{T}\dot{\mathbf{L}} = \mathbf{T}(\mathbf{B}\dot{\mathbf{X}} - \dot{\mathbf{V}}) = -\mathbf{T}\dot{\mathbf{V}} \tag{23.14}$$

The vector $\dot{\mathbf{V}}$ can be calculated by the following.

$$\dot{\mathbf{V}} = (\mathbf{B}(\mathbf{B^T B})^{-1}\mathbf{B^T} - \mathbf{E})\dot{\mathbf{L}} \tag{23.15}$$

Define $\dot{\varepsilon} = -\dot{\mathbf{V}}$, the matrix $\dot{\mathbf{A}}$ is formed is formed as blow.

$$\dot{\mathbf{A}} = [\dot{\varepsilon}_1 \mathbf{T_1} \;\; \dot{\varepsilon}_2 \mathbf{T_2} \;\; \cdots \;\; \dot{\varepsilon}_n \mathbf{T_n} \;\; \dot{\mathbf{t}}] = [\mathbf{a_1} \; \mathbf{a_2} \; \cdots \; \mathbf{a_n} \; \mathbf{a_{n+1}}] \quad (i = 1, 2, \cdots, n) \tag{23.16}$$

The correlation coefficient c_i of the first n vectors and the last vector is calculated as the same way of Eq. (23.13). The unreasonable observation can be eliminated by the value of c_i.

23.2.5 RAIM Flowchart for Rocket Trajectory Determination

The RAIM Flowchart for rocket trajectory determination is show in the Fig. 23.1. It can be observed that the position RAIM and the velocity RAIM are separated. This situation is caused by the measure signal for position and velocity is different. At the same time, the unreasonable pseudo range observation and Doppler frequency observation may not belong a same GPS satellite. The detail flowchart of RAIM position is as following.

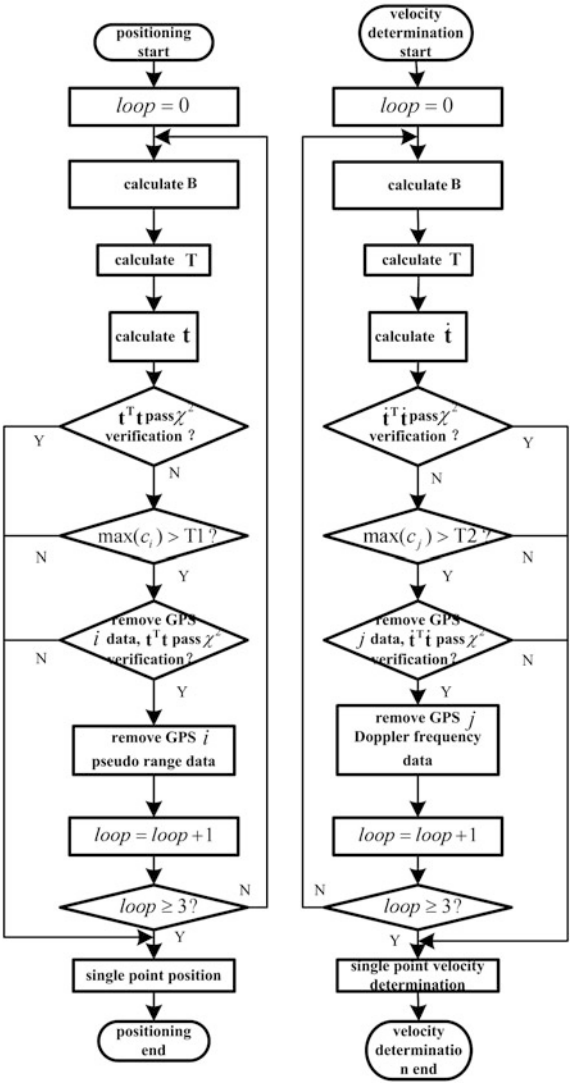


Fig. 23.1 Flowchart of RAIM process in position and velocity measurement

1. *loop* is set to 0, then jump step 2;
2. matrix **B** and the matrix T is calculated, jump step 3;
3. vector **t** is acquired and check the value $\mathbf{t}^{\mathbf{T}}\mathbf{t}$ with χ^2 distribution, if not pass, jump step 4; else jump step 8;
4. judge $\max(c_i) > \mathrm{T1}$, if pass, jump step 5; else jump step 8;
5. remove the observation of the satellite i and calculate $\mathbf{t}^{\mathbf{T}}\mathbf{t}$ again, check the value $\mathbf{t}^{\mathbf{T}}\mathbf{t}$ with χ^2 distribution again, if pass, jump step 6; else jump step 8;
6. remove the observation of the satellite i, form a new equation group, jump step 7;
7. $loop = loop + 1$, judge $loop \geq 3$, if pass, jump step 8; else jump step 2;
8. single point position, end.

Because the rocket onboard GNSS receiver can receive the signal of 8–12 GPS satellites in common, 3 loops are made for RAIM. The multiple loops for RAIM process may eliminate the bad influence caused by the multiple satellites. RAIM for velocity determination is the same as that for position, only the threshold is replaced by T2.

23.3 Verifications

According to the method described in Sect. 2.5, the algorithm is verified with the actual measurement GNSS data in a real launching task. The GNSS data contain the GPS observation data and GPS navigation message in GNSS special format. The data are firstly restored and then made RAIM process for position and velocity determination respectively.

The figures in Fig. 23.2 show the position precision curve with RAIM (marked with blue solid line) and the one without RAIM (marked with red dot line) within 85710–85850 s (Beijing Time) at X, Y and Z directions respectively. It can be observed that the curve with RAIM is more stable within the whole time range and controlled mostly between 0.5 and 1.5 m.

The figures in Fig. 23.3 show the velocity precision curve with RAIM (marked with blue solid line) and the one without RAIM (marked with red dot line) within 85710–85850 s at X, Y and Z directions respectively. It can be observed that the data fluctuation in red dot line are eliminated in the blue solid line and the curve with RAIM is much more stable within the whole time range and controlled mostly between 0.01 and 0.05 m/s.

The method proposed in this paper can remove the unreasonable data adaptively. For an instance in the RAIM for velocity determination, there are no data disable within 85734–85752 s so that two curves in the range are overlapped in Fig. 23.3. The satellite 21 observations are mostly removed within 85763–85821 s and the satellite 25 observations are mostly removed within 85821–85845 s. And after these RAIM process, the velocity precision improves obviously.

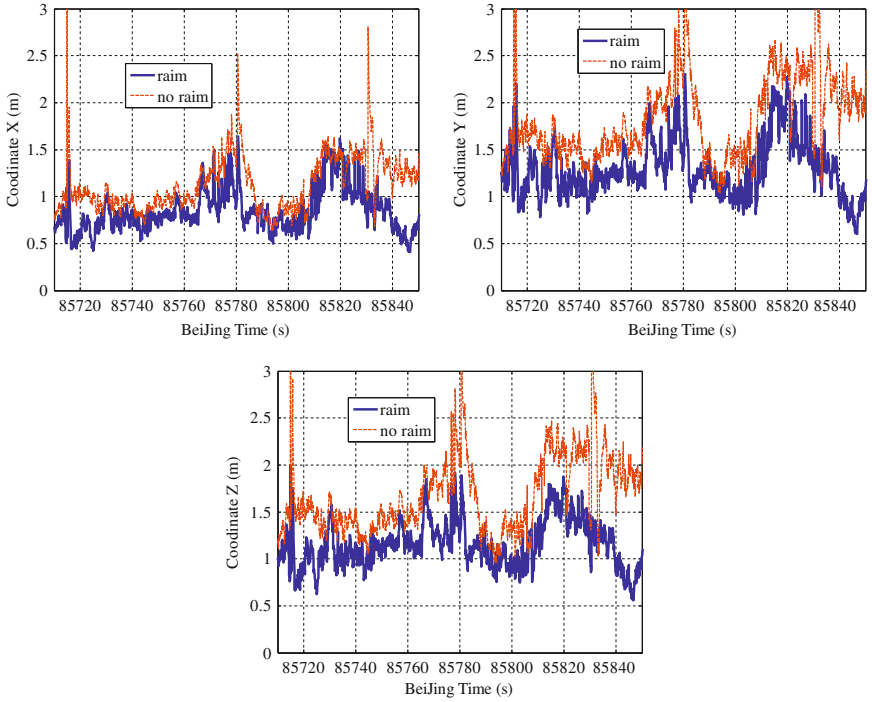


Fig. 23.2 Accuracy curves of position with RAIM

The average precision data of the position and velocity with RAIM and without RAIM are listed in Table 23.1. It can be observed that the average position precision improves about 30 % after RAIM and the average velocity precision after RAIM improves much more obvious, about 90 %.

The rocket onboard GNSS receiver has some characteristics such as high speed and high dynamic, which makes a negative impact on the Doppler frequency data and brings about the excess vibration of the velocity within a great scale (From Fig. 23.3, the red dot line is beyond 0.4 m/s within some ranges), RAIM can remove the unreasonable Doppler frequency data so as to improve the velocity precision by one order of magnitude. Moreover, the high speed and high dynamic environment does not affect the C/A pseudo range data obviously and the original single point position accuracy can be kept at an acceptable level. In this situation, RAIM can also increase the positioning precision about 30 %. In conclusion, RAIM is effective for the high speed and high dynamic GNSS data process, especially in the velocity measurement field.

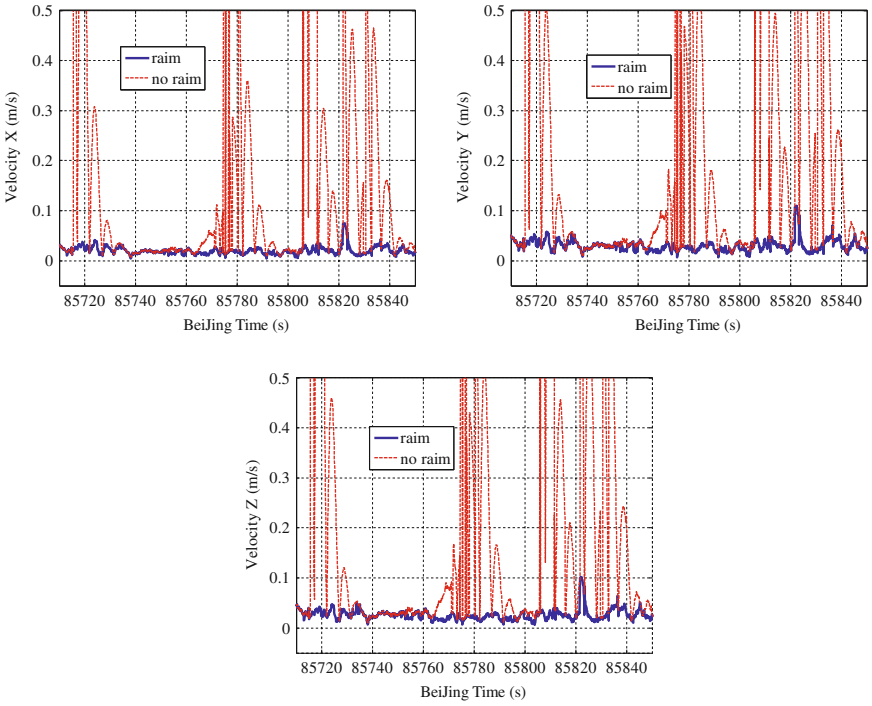


Fig. 23.3 Accuracy curves of velocity with RAIM

Table 23.1 Accuracy of position and velocity

Precision	No RAIM	RAIM	Improvement (%)
X(m)	1.244	0.852	31.5
Y(m)	2.021	1.294	35.9
Z(m)	1.852	1.154	37.7
X(m/s)	0.195	0.020	89.5
Y(m/s)	0.317	0.030	90.5
Z(m/s)	0.292	0.026	91.0

23.4 Conclusions

In this paper, the RAIM method in position field is investigated for the rocket onboard GNSS receiver and then extends to the velocity filed. Moreover, the rocket GNSS data in a real launching task are processed by the proposed RAIM method. The process results can verify that the proposed algorithm can not only improve the position precision effectively, but also remove the velocity fluctuation obviously. The method can be applied in more GNSS data processing fields, such as various missile, rocket and aerocraft.

References

1. Bi J (2010) GNSS integrity monitoring method, technology and application. WuHan University PhD Thesis
2. Ying Li, Meng Li (2012) Study on RAIM algorithm of gps receiver based on parity vector. Geomatics Spat Inf Technol 35(1):158–160
3. Xu L, Li X (2011) A receiver clock bias-based GPS RAIM algorithm. J Astronaut 32(3):537–542
4. Ren K, Li Y, Feng Y (2009) Comparison of two RAIM availability analysis approaches. J Telemetry, Tracking Command 30(2):13–17
5. Wang M, Liu W, Wang F (2009) RAIM availability augmentation algorithm in position and velocity domain aided by precise forecasted time and frequency. J Astronaut 30(3):962–967
6. Zhang S (2004) GPS technology and application. National Defence Industry Press, Beijing

Chapter 24
The Analysis of Character of User Range Accuracy

Yuechen Wang and Rui Li

Abstract The new generation of Global Positioning System (GPS) plans to use user range accuracy (URA) as the integrity monitoring parameter of the signal in space. In the future, the system can forecast and monitor user range error (URE), which results from the signal-in-space error (SISE), including ephemeris error and clock error. By this way, system can reduce the space segment incidents' impact for navigation users effectively. URA is an original parameter of GPS's message, but its definition and application are changing all the time. In this paper, we firstly summary various changes of the URA, include not only the definition and the application, but also the refining of the index. Then, for the sake of finding the reasons of these changes, we make statistics on the IGS's broadcast and precise ephemeris from 2002 to 2012. By analyzing the character of SISE and URE, we get conclusions below: because of the accuracy's improvement of ephemeris and clock, the accuracy of URA must be improved. What's more, the URE's mean is non-zero, by the using of the concept of "bound", the using of original URA will lead to an exaggerated bound so that it can't reflect the integrity of the system. We can use the concept of error ellipsoid to explain this phenomenon effectively.

Keywords GPS · User range accuracy · Integrity monitoring · Error ellipsoid

24.1 Introduction

GPS is a satellite navigation system of USA. So far, it is widely used in the field of measurement, navigation, meteorology and so on. Besides the improvement of the system's accuracy, the real-time, continuity and integrity abilities have attracted

Y. Wang (✉) · R. Li
School of Electronic and Information Engineering of Beihang University,
Beijing 100191, China
e-mail: wangyuechenpavel@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_24,

more and more attentions. Especially, in the context of the civil aviation applications, the system's integrity performance is particularly important. The integrity means, when the navigation system can't be used, it must have the ability of alarming users in time [1, 2]. The integrity monitoring includes three important parameters: integrity risk, alarm time and alarm threshold [2].

The new generation of GPS system plans to use user range accuracy (URA) as the integrity monitoring parameter of the signal in space. In order to achieve this goal, GPS has carried out a series of changes on URA. In this paper, we firstly search the literature and list the changes of the definition and applications of URA. Then, according to the statistics of data, we analyze the reasons of all changes by using the error ellipsoid, which can show the significance of the changes for integrity monitoring. All of the studies will contribute to a better understanding of the application of URA in integrity monitoring.

24.2 User Range Accuracy

URA is an original GPS's message parameter and has been changed lots times, including definition, accuracy and parameters' separation as the following.

24.2.1 The Definition of URA

In the GPS's ICD-GPS-200C in 1997, URA is defined as the 1σ estimation of the user range error (URE). URE contains errors caused by satellite's ephemeris and clock, but doesn't include the signal transmission and receiver errors [3]. So URA is only an estimation of URE's range and can't be used to monitor integrity. Then, in 2004, the IS-GPS-200D made no change on URA [4].

In 2010, IS-GPS-200E first links URA to integrity monitoring. Although it is still defined URA as the 1σestimation of URE, with the appearance of Integrity Status Flag (ISF), URA is divided into two kinds of accuracy. One is legacy level, the system must satisfy the integrity risk of $10^{-5}/h$, which means the incident's undetected probability must be lower than $10^{-5}/h$ by the threshold of 4.42 times URA. Another is enhanced level, the ISF is set "on" and the system must satisfy the integrity risk probability of $10^{-8}/h$ [5]. By these additions, URA has the ability of integrity monitoring.

The latest specification IS-GPS-200F was released in 2011. Now URA's definition is modified as the conservative root-mean-square (RMS) estimation of URE. It retains 200E's "ISF" concept and lists several situations that satellites are unavailable, such as an alert is issued to users before the URE exceeds either of the scaled URA bounds, or an alert is issued to users no more than 8.0 s after the URE exceeds the 4.42 times URA bound, or if the integrity status flag is 'on' and an

alert is issued to the users no more than 5.2 s after URE exceeds the 5.73 times URA bound. The system's "alert" includes broadcasting satellite's unhealthy message, non-standard code or parity error, etc. [6]. These methods can alarm users that the satellite is unavailable for navigation so that the system's integrity can be insured.

24.2.2 The Index of URA

URA is broadcasted in the GPS's message. Due to the limit of massage's capacity, broadcasted URA is not its value, but an index value. Each index corresponds to a specific URA range. The original URA reflects the approximate range of URE and is divided into 0 to 15, as be shown in Table 24.1 [3].

Starting from the 200D, the 0 index is subdivided into -16–0, as be shown in Table 24.2 (Note:1–15 indexes are the same as origin and doesn't show here). As can be seen from Table 24.2, URA's accuracy improves from 2.40 to 0.01 m [4]. In the future, each URA index's bound will be used as the integrity monitoring value.

24.2.3 The Parameter's Decomposition of URA

What's more, URA is divided into various parameters. In IS-GPS-200D, they are URA_{oe} and URA_{oc}, respectively presents ephemeris and clock errors. In addition, URA_{oc} is divided into three parameters:URA_{ocb}, URA_{oc1} and URA_{oc2} [4]. In IS-GPS-200F, URA is divided into URA_{ED} and URA_{NED}, respectively presents the elevation-dependent (ED) and the non-elevation-dependent (NED) errors. URA_{ED} mainly reflects the along-track and cross-track ephemeris error, while URA_{NED} primarily reflects radial ephemeris error and clock error. In the same way, URA_{NED} is composed by URA_{NED0}, URA_{NED1} and URA_{NED2} [6].

Table 24.1 The index of the original URA

URA index	URA (m)	URA index	URA (m)
0	$0.00 < URA \leq 2.40$	8	$48.00 < URA \leq 96.00$
1	$2.40 < URA \leq 3.40$	9	$96.00 < URA \leq 192.00$
2	$3.40 < URA \leq 4.85$	10	$192.00 < URA \leq 384.00$
3	$4.85 < URA \leq 6.85$	11	$384.00 < URA \leq 768.00$
4	$6.85 < URA \leq 9.65$	12	$768.00 < URA \leq 1536.00$
5	$9.65 < URA \leq 13.65$	13	$1536.00 < URA \leq 3072.00$
6	$13.65 < URA \leq 24.00$	14	$3072.00 < URA \leq 6144.00$
7	$24.00 < URA \leq 48.00$	15	$6144.00 < URA$ (or no accuracy prediction is available)

Table 24.2 The index of the URA now

URA index	URA (m)	URA index	URA (m)
0	$1.70 < URA \leq 2.40$	−8	$0.11 < URA \leq 0.15$
−1	$1.20 < URA \leq 1.70$	−9	$0.08 < URA \leq 0.11$
−2	$0.85 < URA \leq 1.20$	−10	$0.06 < URA \leq 0.08$
−3	$0.60 < URA \leq 0.85$	−11	$0.04 < URA \leq 0.06$
−4	$0.43 < URA \leq 0.60$	−12	$0.03 < URA \leq 0.04$
−5	$0.30 < URA \leq 0.43$	−13	$0.02 < URA \leq 0.03$
−6	$0.21 < URA \leq 0.30$	−14	$0.01 < URA \leq 0.02$
−7	$0.15 < URA \leq 0.21$	−15	$URA \leq 0.01$
		−16	No accuracy prediction available-use at own risk

24.3 Data Calculate and Statistical Characterization

In order to know the purposes and reasons of URA's changes, we make statistics.

24.3.1 The Acquisition and Process of Data

As URE is mainly caused by the SISE, we select IGS's broadcast and precise ephemeris from 2002 to 2012 to make statistics, the broadcast ephemeris is brdc's file of ".n" and the precise ephemeris is theigs's file of ".sp3". Mainly including three steps:

- The calculation of broadcast ephemeris: extracting satellite's orbital and clock parameters from broadcast ephemeris. According to the interface specification, calculating satellite's position in the Earth-centered Earth-Fixed (ECEF) coordinate system and clock error every 900 s.
- The calculation of SISE: extracting GPS precise ephemeris of the same day and getting precise orbit and clock data. Using Lagrange interpolation to calculate errors of the same moment and doing subtraction, getting the satellite's ephemeris and clock errors in the ECEF coordinate. Then, represented errors in the reference frame with respect to the satellite by the rotation of coordinate system, just as radial, along-track and cross-track errors.
- The calculation of URE: through the steps above, we get the ephemeris and clock errors. Then using the following formula (24.1) [7] to calculate URE.

$$URE = \sqrt{(0.98R - T)^2 + 0.141^2(A^2 + C^2)} \tag{24.1}$$

In this formula, R, A and C respectively represents ephemeris' radial, along-track and cross-track error, T is satellite's clock error, their units are all Meter.

Using statistical methods, we can get mean and variance of SISE and URE, which can tell us the changes of system by the comparison between different satellites, different years.

24.3.2 The Statistics and Analysis of Data

By the analysis of numerical characteristics, we get the characteristics of SISE and URE, summarize regularities as the following.

24.3.2.1 The statistical character of SISE

Because of the limited space, we select 2002, 2007 and 2012's data to show the changes, years which aren't shown here also comply with the law. We classify satellites by the satellite number (SVN) and category (IIA, IIR, IIR-M and IIF), calculate ephemeris' radial (error_R), cross-track (error_C), along-track (error_A) and clock (error_T) errors, showing their mean and variance in Figs. 24.1, 24.2, 24.3. The red point represents mean and the blue line represents variance.

Through the results, we conclude various regulations of SISE. Firstly, the accuracy of SISE improves: the ephemeris' accuracy improves from 10 m in 2002 to less than 5 m in 2012 and the clock's changes from 10–20 m to 5 m. Secondly,

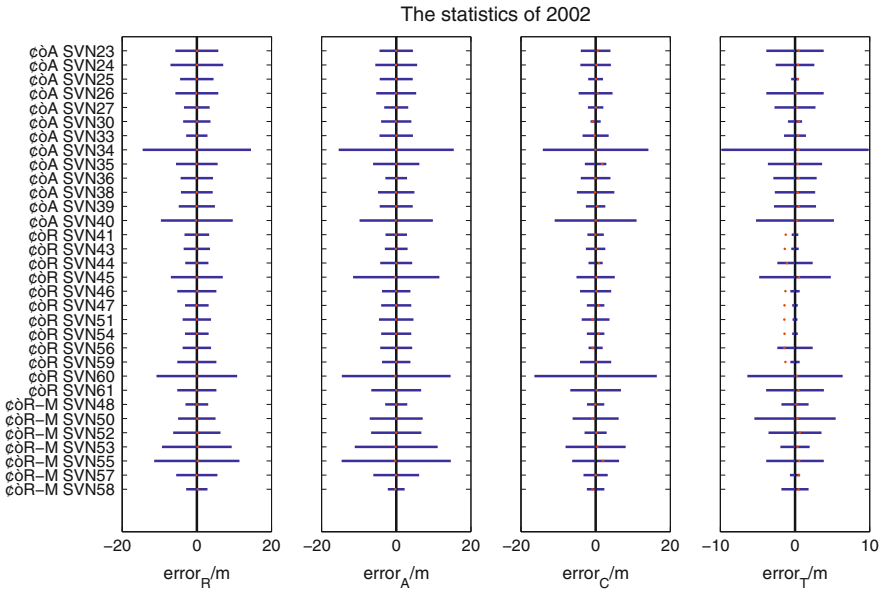


Fig. 24.1 The error statistics of 2002

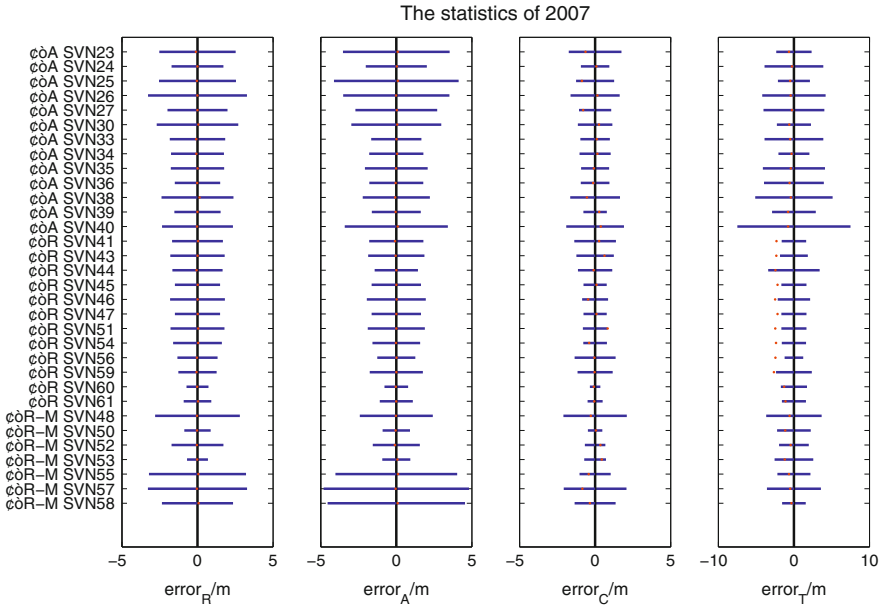


Fig. 24.2 The error statistics of 2007

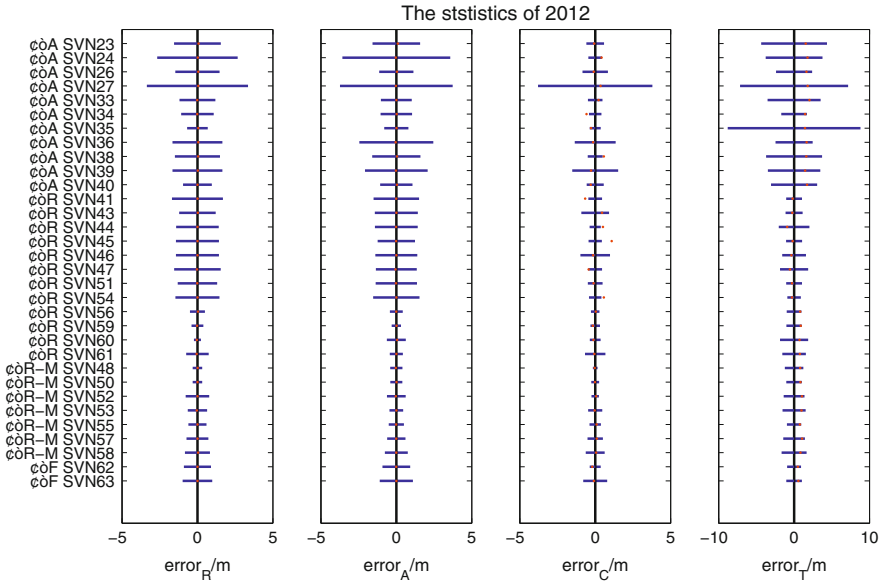


Fig. 24.3 The error statistics of 2012

the accuracy of SISE is related to satellite's category: the new satellites are much better than the olds, both ephemeris and clock's accuracy are less than 2 m.

The accuracy of SISE is improving and ephemeris, clock error has been nearly the same magnitude, so monitoring ephemeris and clock respectively by $URA_{oe}$and URA_{oc}will be beneficial to find and solve incidents. Since the satellite clock error varies quickly, more parameters can describe clock error accurately. In addition, considering the clock error mainly concentrates on the radial direction [7], we can merge the clock and radial ephemeris errors into URA_{NED} so that we can separate URA into URA_{ED} and URA_{NED}. But the two separations are similar in affection.

24.3.2.2 The Statistical Character of URE

In this part, we still select data of 2002, 2007 and 2012 to analyze URE's changes. URE's mean and variance are shown in the Fig. 24.4, arranging the graphs from left to right chronologically, red dot represents URE's mean and the blue line is the variance.

From the figure, we can get the following conclusions. Firstly, with the improvement of SISE's accuracy, URE's accuracy is also improving: the mean reduces from about 3 m to less than 2 m and the variance changes from about 5 m to 1–2 m, which means URE's mean and variance are substantially equal. For the original URA index's accuracy can only accurate to 2.4 m, it isn't sufficient enough to reflect URE's change. So the detail of index is necessary. Secondly,

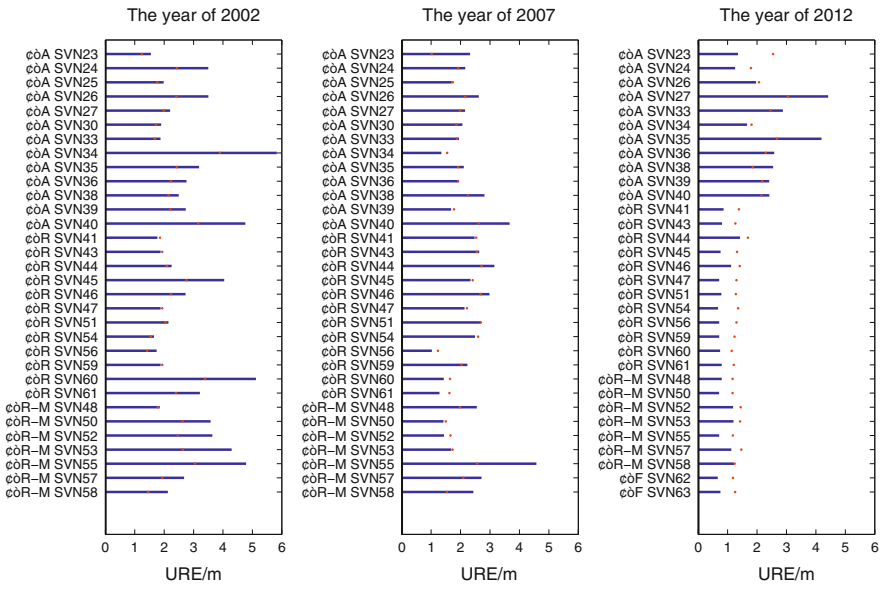


Fig. 24.4 The error statistics and comparison of URE

URE's mean is non-zero because all satellites' means are 1–2 m. With the using of the concept of "bound" for integrity monitoring, using the original URA will lead to an exaggerated bound.

The concept of "bound" is special: while estimating URA, system wants to use a zero-mean Gaussian distribution to describe the real URE's distribution. The schematic diagram of probability density function is shown in Fig. 24.5. In the figure, the blue line is a distribution of N(1.2, 0.8), as the URE's distribution, the red line stands for a distribution of N(0, 1.5), as a zero-mean Gaussian.

What needs to be paid more attention is that "bound" mainly requires the censored probability: by the threshold, if the zero-mean Gaussian distribution's censored probability is less than the integrity risk, URE's censored probability can satisfy the integrity risk surely, just as the blue and red area show in Fig. 24.5: the red is larger than the blue significantly.

Now, let's analyze the reasons why original URA will cause the bound exaggerated. Defined as URE's 1σ estimation, original URA can be seen as URE's 68 % bound. Setting URE's mean $\mu \geq 0$, variance $\sigma \geq 0$, so

$$URA = \mu + \sigma \tag{24.2}$$

When using it in the integrity monitoring, in order to satisfy the integrity risk of $10^{-5}/h$, the bound should be 4.42 times URA

$$URE_{bound} = K \cdot (\mu + \sigma) \tag{24.3}$$

In the new definition of URA, it's URE's RMS estimation. Taking URA_{NED} as example, it includes radial and clock errors. Considering the clock error mainly concentrated in the radial direction, merging them into a non-zero mean error. During integrity monitoring, corresponding to integrity risk of $10^{-5}/h$, expanding the variance in 4.42 times and adding to the mean. The bound is

$$URE'_{bound} = \mu + K \cdot \sigma \tag{24.4}$$

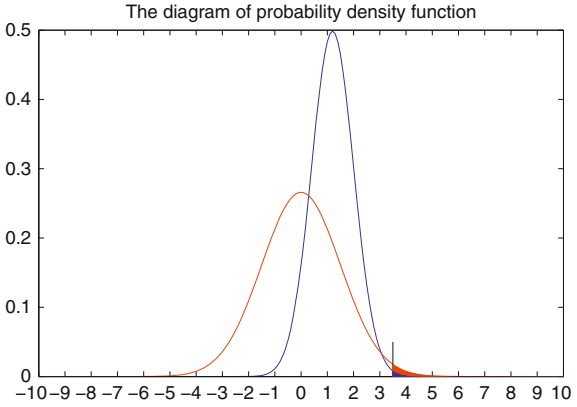


Fig. 24.5 The schematic diagram of probability density function for the bound

The K is Gaussian distribution quantile. Here, K = 4.42. So the new URA is:

$$URA' = \frac{1}{K} URE_{bound} = \frac{1}{K}\mu + \sigma \tag{24.5}$$

Comparing the (24.2) and (24.5), obviously

$$URA \geq URA' \tag{24.6}$$

If and only if $\mu = 0$, equation holds. But statistics show that URE' mean is nonzero, so URA in the new definition is always smaller than the original one.

While, URA_{ED} includes the along-track and cross-track errors of ephemeris and can't be merged. So using error ellipse to analyze, the same conclusions can be obtained as above [8]. The new URA's calculation is related to the concept of "bound", so it's different from the origin. In one word, the new URA is smaller and more beneficial for the integrity monitoring.

24.4 The Verification by Statistics of Data

In this part, we use IGS's data to verify whether URA's bound is exaggerated. Now, the broadcast ephemeris' URA is still the old and isn't separated, so we can compare 4.42 times URA with the calculated URE.

We select the data in 2012, from April 1 to June 30. Showing one satellite's result each category as example in Fig. 24.6, they are SVN34 (IIA), SVN41 (IIR),

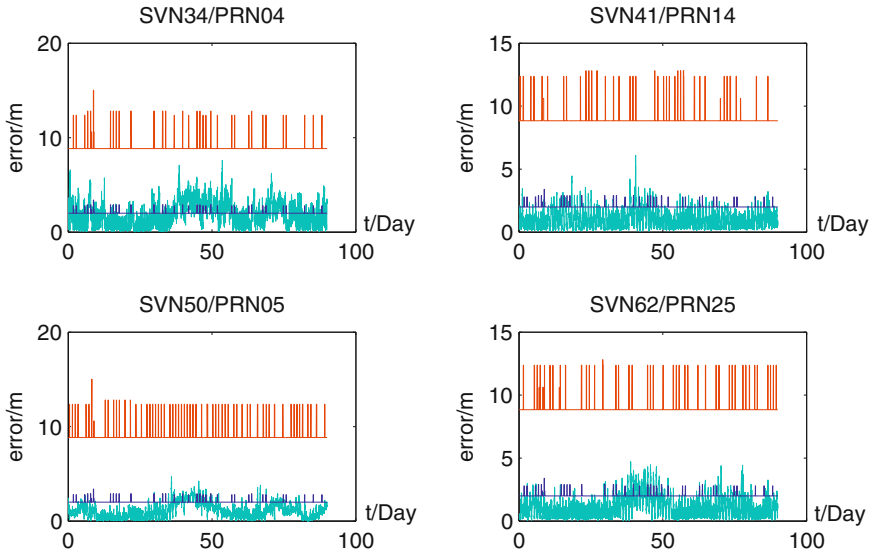


Fig. 24.6 The comparison between URA and URE

SVN50 (IIR-M) and SVN62 (IIF). URE is the green and URA is blue, red line stand for 4.42 times URA.

It can be seen from Fig. 24.6 that most of the broadcast URA is 0 or 1, corresponding to 2 or 2.8 m's URE, but the URE's value is always 1–2 m and may be less than 1 m sometimes. URA's index can't reflect the status of URE accurately.

What's more important, the bound of 4.42 times URA exaggerates URE and it will be serious with the update of the category. The bound can't reflect URE's magnitude and trend. So URA's exaggerated bound can't satisfy the requirement of integrity now.

These two conclusions verify that the previous theoretical analysis is correct. It not only shows the necessity of the changes on current URA, but also tells us that the error ellipsoid can bound URE feasibly and correctly.

24.5 Summary

Statistics and analysis of the above data, getting the following conclusions:

1. Satellite's SISE's accuracy improves from 10 m of ephemeris and 20 m of clock a decade ago to 5 m both of them. It results in the parameter separation of URA;
2. The improving of SISE's accuracy leads to the improvement of URE's accuracy. URE's mean and variance are both approximately 2 m so that the original URA index accuracy can't reflect the URE sufficiently. So the detailing of URA index is necessary;
3. URE's mean is non-zero and when using the original URA to monitor integrity, the bound will be exaggerated and can't meet the required integrity risk. So the change of URA's definition is imperative;
4. The error ellipsoid model can effectively explain reasons of the exaggerated bound. It may be an effective way to calculate URA.

Acknowledgments The authors gratefully acknowledge the support of the National Basic Research Program of China (Grant No. 2010CB731805).

References

1. Kaplan ED, Hegarty CJ (2010) Understanding GPS: principles and applications, 2nd Edn[M]. Kou Yan-hong, translation. Publishing House of Electronics Industry, Beijing, 3:256–258
2. Fan M, Zhou J, Niu F, Chang Z (2011) Analysis for satellite navigation system basic integrity algorithm and performance. J Geomatics Sci Technol 6
3. GPS interface control document ICD-GPS-200C. http://www.navcen.uscg.gov

4. GPS interface specification IS-GPS-200D
5. GPS interface specification IS-GPS-200E
6. GPS interface specification IS-GPS-200F
7. Heng L, Gao GX, Walter T, Enge P (2011) Statistical characterization of GPS signal-in-space errors [C]. ION National Technical Meeting 2011, San Diego, California, Jan
8. Shao B, Liu J, Zhao R, Huang Z, Rui L (2011) A user differential range error calculating algorithm based on analytic method. Chin J Aeronaut 24:762–767

Chapter 25
Ionosphere Integrity Monitoring on the BDS

Chuanhua Zhao, Jinzhong Bei and Yamin Dang

Abstract Integrity is the ability which the system sends out warnings to users when the navigation system has any failures or outliers overrunning, it can't be used for navigation and positioning. Given the user safety, the integrity of the navigation system is more important than the precision. The BDS is a new GNSS system following GPS and GLONASS, which is carried out by China independently. In order to better the global service of the BDS, it is necessary to induct integrity monitoring with the on the BDS. Ionosphere delay is one of the biggest error sources after the United States to canceled SA policy. Its changes are the most complex and difficult to model accurately. Ionosphere integrity monitoring must be conducted in order to guarantee the reliability and availability of positioning by mobile users. This paper is based on the BDS data of CORS network in Hebei province to conduct ionosphere integrity monitoring. The results show that the ionosphere delay calculated by these data is stable, and the data are available. Two integrity monitoring indicator IRIM and IRIU both reflect the situation of regional ionosphere delay. It has important significance to study the BDS.

Keywords BDS · Integrity monitoring · Ionosphere delay · IRIM · IRIU

25.1 Introduction

COMPASS satellite navigation system is a global satellite navigation system of self-development and independent running which is being carried out by China. According to the overall plan of the system construction, 2012 years or so, the

C. Zhao (✉)
Shandong University of Science and Technology, Qingdao 266510, China
e-mail: zhaochuanhua2007@126.com

C. Zhao · J. Bei · Y. Dang
Chinese Academy of Surveying and Mapping, Beijing 100830, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_25,

system will firstly have the ability of positioning, navigation, timing services and short message communication service in the Asia-pacific region. Around 2020, the worldwide COMPASS system will be built. Integrity is the ability which the system sends out warnings to users when the navigation system has any failure or error overrunning, and it can't be used for navigation and positioning. The integrity information of the system is required no matter when and where, and they are important parameters which guarantee the safety of users [1]. Ionosphere delay error is one of the biggest error sources after the United States canceled SA policy. Its changes are the most complex, so it is significant to study the integrity of the ionosphere.

This paper puts forward ionosphere integrity monitoring of the COMPASS system based on the Hebei province CORS network, and has monitored the measured data. Some useful conclusions are obtained.

25.2 Algorithm Processing

The ionosphere integrity monitoring of COMPASS system includes the following four aspects. Firstly, calculation and analysis of ionosphere delay; Secondly, construction of ionosphere interpolation model; Thirdly, calculation of ionosphere residual integrity monitoring; Fourthly, calculation of the ionosphere residual interpolation uncertainty.

25.2.1 The Calculation of Ionosphere Delay

Since the ionosphere delay has dispersiveness, it can be calculated using the geometry non-combination formed by double-frequency carrier phase. After the ambiguity is fixed, and multi-path effect is fully weakened, the double-difference ionosphere delay can be calculated accurately. The time and space change rule of ionosphere delay can be analyzed quantitatively (such as the relationship with local time, baseline length, satellite altitude angle, and the reference station geographical location) [2]. When the ambiguity is fixed, the double-difference carrier phase observation noise will become very small, so the observation noise can be ignored and cm-level accuracy of the ionosphere delay can be extracted [3].

$$\Delta\nabla I = \left(\frac{f_2^2}{f_1^2 - f_2^2}\right)[(\lambda_1\Delta\nabla\varphi_1 - \lambda_2\Delta\nabla\varphi_2) + (\lambda_1\Delta\nabla N_1 - \lambda_2\Delta\nabla N_2)]s \qquad (25.1)$$

Now the measured data of COMPASS is double frequency, so the above type also applies to COMPASS system. In this paper, double difference ionosphere delay of COMPASS system is calculated using above type.

25.2.2 Establishing the Ionosphere Interpolation Model

Various interpolation models include linear combination model, low-order surface model, linear interpolation model, distance based linear interpolation model and least-square collocation model. Various interpolation models don't have significant difference in the interpolation spatial correlation error aspects according to the study of Dai in 2001 [4]. Because low-order surface model has very good approximation and fitting within the scope of the ionosphere delay spatial correlation [5], so the weighted low-order surface model is used in the calculation of double-difference ionosphere delay interpolation.

Mathematical model is the following [6]:

$$\Delta\nabla I_{i,n} = \alpha_1 \Delta X_{i,n} + \alpha_2 \Delta Y_{i,n} + \alpha_3 \tag{25.2}$$

When the reference station number (given n) is larger 3, the least squares solution of interpolation coefficient vector can be obtained.

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \left(A^T P A\right)^{-1} A^T P L,\ A = \begin{bmatrix} \Delta X_{1,n} & \Delta Y_{1,n} & 1 \\ \Delta X_{2,n} & \Delta Y_{2,n} & 1 \\ \vdots & \vdots & \vdots \\ \Delta X_{n-1,n} & \Delta Y_{n-1,n} & 1 \end{bmatrix}$$
$$L = \begin{bmatrix} \Delta\nabla I_{1,n} \\ \Delta\nabla I_{2,n} \\ \vdots \\ \Delta\nabla I_{n-1,n} \end{bmatrix},\ P = diag\left(d_{u,1}^{-1}, \ldots, d_{u,n}^{-1}\right) \tag{25.3}$$

where, d is the distance between rover station and every reference station. The double-difference ionosphere delay between rover station and the main reference station can be represented as following.

$$\Delta\nabla I_{u,n} = B \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{25.4}$$

where, $B = [\,\Delta X_{u,n} \quad \Delta Y_{u,n} \quad 1\,]$ ΔX and ΔY are the coordinate difference in the same coordinate system respectively.

25.2.3 Ionosphere Residual Integrity Monitoring Indicator

Chen X. put forward two non-difference ionosphere linear index in 2003, they are ionosphere residual integrity monitoring indicator and ionosphere residual

interpolation uncertainty indicator. Japan and Germany's network data are used in the experimental analysis, and the results show that the indicator can better estimate ionosphere interpolation error and ionosphere residual error of the rover station [7].

The following is the calculation process of IRIM. After calculating the double-difference ionosphere delay of baseline by type (1) (namely calculated value), then using double-difference ionosphere delay of several baselines (at least two) to interpolate ionosphere delay of another baseline (namely estimator). Then the difference between the calculation and interpolated estimator is compared. All satellite weights (weighted method for high angle weighted method) root mean square (RMS) of each epoch is calculated, and then the weighted RMS is cumulated (cumulative interval takes an hour). The weighted method is calculated as follows.

$$P(z) = cos^2(z) \tag{25.5}$$

where, z is the high angle of satellites in reference stations.

25.2.4 Ionosphere Residual Interpolation Uncertainty

Although IRIM can reflect the overall situation of the ionosphere, but the interpolation uncertainty of ionosphere delay in every epoch still need to consider for some users. Interpolation uncertainty can reflect the reliability of positioning. IRIU is used for this purpose.

The calculation principle of IRIU is following. Firstly, through a few reference stations (at least 4) around the rover station, the regional interpolation model is constructed by several baselines. Secondly interpolation standard deviation is calculated. The standard deviation represents the linear degree of ionosphere interpolation in the client area. Thirdly the weight average value of the interpolation standard deviation of all satellites with some moment is calculated. IRIU is the weighted average value [8].

Interpolation standard deviation is calculated as follows.

Firstly, interpolation coefficient α is calculated.

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \left(A^T P A\right)^{-1} A^T P L \tag{25.6}$$

Secondly, the unit weight error of α is calculated.

$$\delta_0^2 = \frac{V^T P V}{n-3}, V = \begin{bmatrix} \Delta\nabla V_{1,n} \\ \Delta\nabla V_{2,n} \\ \vdots \\ \Delta\nabla V_{n-1,n} \end{bmatrix}, P = diag\left(d_{u,1}^{-1}, \ldots, d_{u,n}^{-1}\right) \tag{25.7}$$

where, n is the number of reference stations, d is the distance between rover station and reference stations.

Thirdly, the covariance matrix of α is calculated.

$$Q = \delta_0^2 \cdot \left(A^T P A\right)^{-1} \tag{25.8}$$

Fourthly, the variance of $\Delta\nabla I_{u,n}$ is calculated.

$$\delta_{\nabla\Delta I}^2 = BQB^T \tag{25.9}$$

Through the comparison between the double-difference ionosphere delay (absolute value) and its standard deviation, it can test the ionosphere interpolation uncertainty, and the interpolation standard deviation can be used for ionosphere interpolation uncertainty. IRIU is the weighting average value of interpolation standard deviation.

25.3 Example Analyses

In this paper the CORS network data of Hebei province is used and the time form 1:00:00–23:59:30 in August 2010. The sampling rate is 30 s. LC, LX, MC, NP, and SZ five stations are the research objects. SZ station is the main reference station, and NP station is rover station. Figure 25.1 is the statistics of visible satellite numbers for five stations. From Fig. 25.1 we can see that visible satellite numbers are form six to ten. The distribution is uniform.

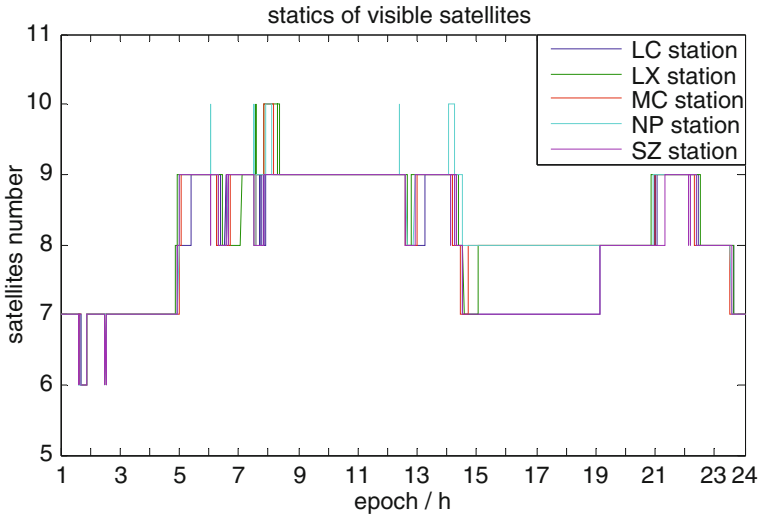


Fig. 25.1 Statistics of visible satellites number

25.3.1 Calculation of the Ionosphere Delay

From Fig. 25.2 we can see that all of the delay value less than 1 m and achieve the accuracy of decimeter level. The ionosphere delay is larger because the four baselines are longer. The ionosphere delay value between 3 and 7 PM is bigger compared to other times; the main reason is that Hebei is located in the east eight time zones. GPS time 3–7 response to from 11 to 15 of local time, at this time the sun radiation is more violent, and ionosphere activity is more severe, so the ionosphere delay is larger. When GPS time 17–2 responses to from 1 to 10 of local time, the section is mainly night and morning, the sun radiation weak and layer is stable, so the ionosphere delay value is smaller.

25.3.2 Ionosphere Interpolation Model

The delay estimator of SZ–NP baseline is interpolated by other three baselines' delay in this paper. The interpolated residuals are shown in Fig. 25.3. Interpolated residuals achieve decimeter level can be seen form Fig. 25.3, but most of them are in the cm-level accuracy. So this method can be applied to ionosphere delay interpolation and mode precisely.

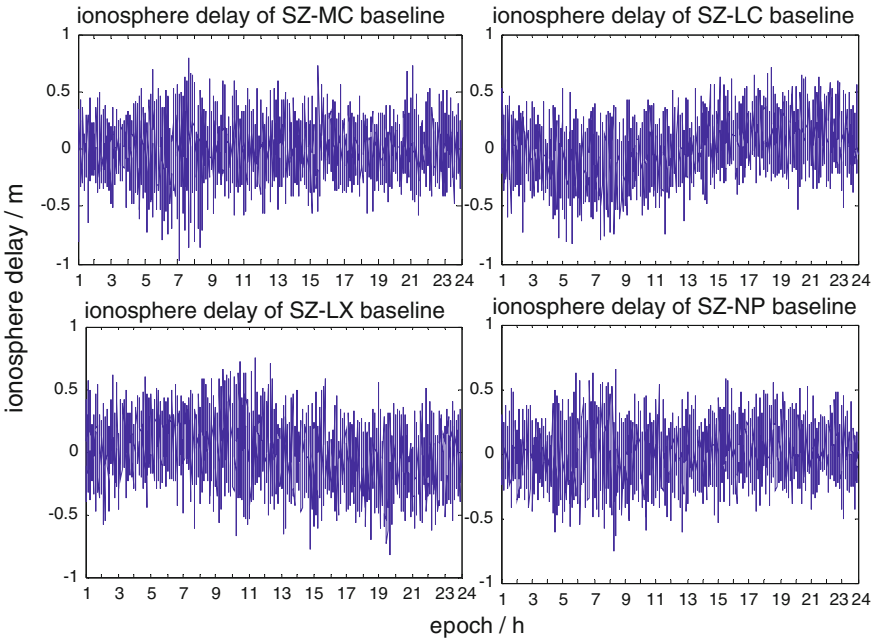


Fig. 25.2 Ionosphere delay of four baselines

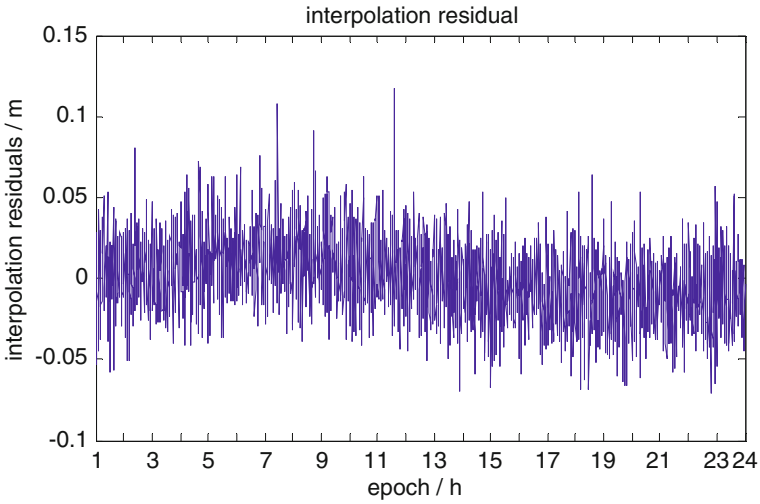


Fig. 25.3 Interpolation residuals

25.3.3 Calculation of Ionosphere Residual Integrity Monitoring Indicator

Figure 25.4 is the IRIM value of SZ–NP baseline. Form Fig. 25.4 we can see that the overall change trend of IRIM value is relatively flat. IRIM values are bigger from 5 to 7, which correspond from 13 to 15 of local time, the solar activity violent, and

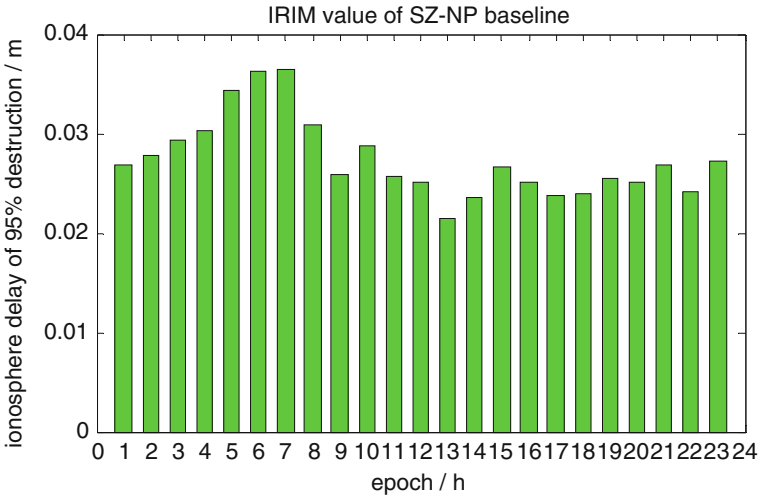


Fig. 25.4 IRIM value of SZ–NP baseline

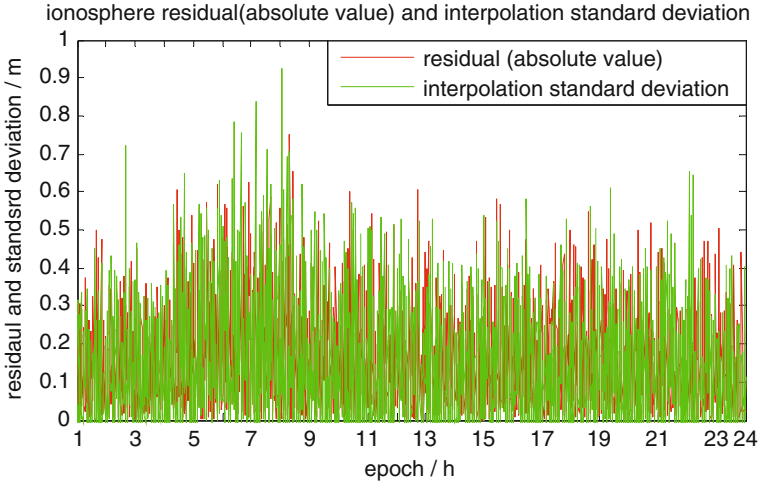


Fig. 25.5 Ionosphere residual (absolute value) and interpolation standard deviation

radiation is stronger so the ionosphere delay is bigger also. The change trend of IRIM and the trend of double-difference ionosphere delay has the very good consistency. It fully illustrates that IRIM monitor the ionosphere residuals effectively. From the indicator we can see the overall situation of ionosphere delay.

25.3.4 Calculation of the Ionosphere Residual Interpolation Uncertainty Indicator

From Fig. 25.5 we can see that the absolute value of ionosphere delay residual and interpolation standard deviation fits better. The trend is the same. So interpolation standard deviation can be used to show ionosphere interpolation uncertainty.

Figures 25.6 and 25.7 are IRIU value of at 08:45 and 20:10 respectively. Two figures reflect the distribution of double-difference ionosphere delay residual in space intuitively. Compared to Fig. 25.6, the IRIU of Fig. 25.7 is relatively larger, because the ionosphere residuals at 20:10 moment are larger than at 08:45. Its interpolation accuracy is smaller. But overall, IRIU correctly reflects ionosphere interpolation situation in this area.

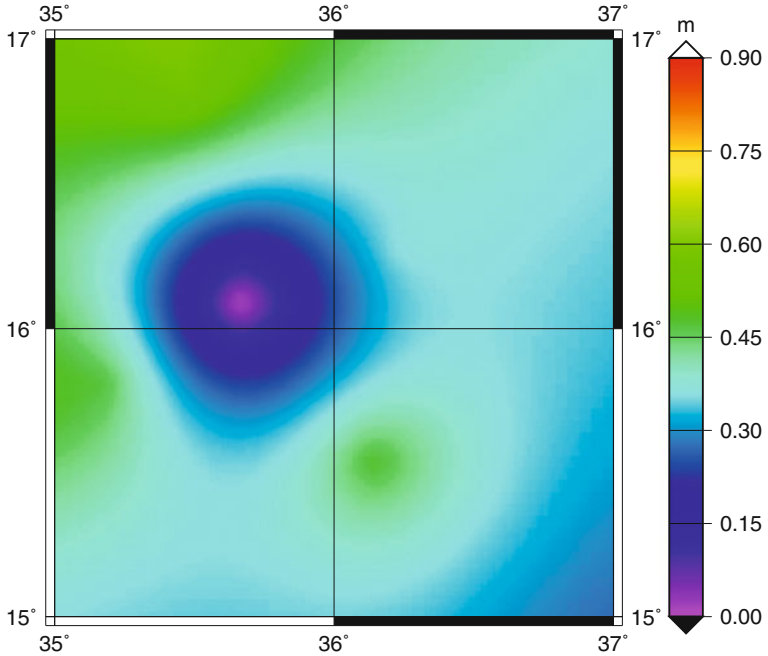


Fig. 25.6 IRIU value at 08:45 moment

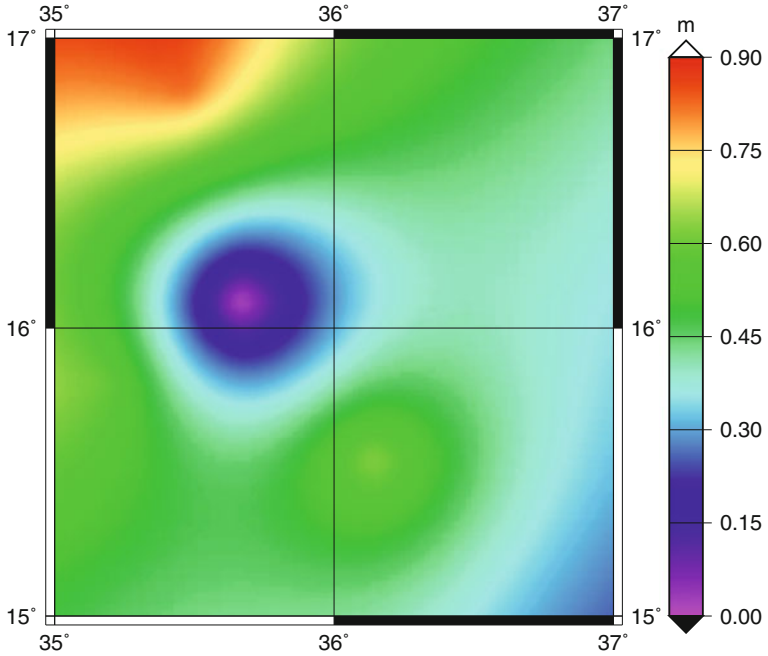


Fig. 25.7 IRIU value at 20:10 moment

25.4 Conclusions

Through the calculation and analysis of double-difference ionosphere delay with four baselines, the conclusion can be obtained that most delay value has reached decimeter level accuracy. The ionosphere residual integrity monitoring indicator can monitor the ionosphere residuals effectively and ionosphere residual interpolation uncertainty indicator reflects the distribution of double-difference ionosphere residual range in space area with a moment. Two indicators can reflect the ionosphere delay and interpolation situation. The ionosphere of above layer is in a more stable state. Because of the limitation of the measured data, the existing data are measured nearly 2 years, the sun activity is stable. The monitoring analysis of the ionosphere abnormalities activities (such as solar maximum) will be the research focus in the next step.

Acknowledgments This research is funded by Shandong University of Science and Technology Graduate Student Science and Technology Innovation Fund (No:YCA120205), the National '863 High Technology Project' (No:2011AA120503), the 'Technology Project' (No:2012BAB16B01), the CASM Research Foundation (No:7771120) and Foundation (No:201103).

References

1. Bei J (2010) GNSS integrity monitoring method, technology and application. Wuhan University, Wuhan, pp 18–19
2. Bai M (2009) Study on precision analysis and integrity of the QDCORS of the network. Shan dong University of Science and Technology, Qingdao, pp 44–45
3. Qiu L, Chen Y, Duan Y (2010) Research on ionosphere delay modelling technology based on GPS RTK network. J Geodesy Geodyn 2(30):57–58
4. Dai L, Wang J, Rizos C, Han S (2003) Predicting atmospheric biases for real-time ambiguity resolution in GPS/GLonass reference stations networks. Accepted by, J Geodesy, 2003
5. Dingfa H, Letao Z, Chenggang L (2009) The theory of GPS augmentation reference station network. Science Press, Beijing, pp 197–199
6. Zhang W (2008) Construction and research of virtual reference station. Jilin University, Jilin
7. Chen X, Landau H, Vollath U (2003) New tools for network RTK integrity monitoring. In: Proceedings of ION GPS/GNSS, Protland, OR, pp 1355–1360
8. Zhou D (2007) Investigation into quality control and integrity monitoring for VRS technique. Southwest Jiao tong University, Sichuan

Part III
Satellite Navigation Models and Methods

Chapter 26
GNSS Integer Ambiguity Estimation and Evaluation: LAMBDA and Ps-LAMBDA

Bofeng Li, Sandra Verhagen and Peter J.G. Teunissen

Abstract Successful integer carrier-phase ambiguity resolution is crucial for high precision GNSS applications. It includes both integer estimation and evaluation. For integer estimation, the LAMBDA method has been applied in a wide variety of GNSS applications. However, before conducting ambiguity resolution, one needs to infer how reliable the fixed solution is expected to be, as incorrect fixed ambiguity solutions often lead to unacceptable positioning errors. In this paper, two Matlab software tools are introduced for the evaluation and integer estimation: Ps-LAMBDA and an updated version of LAMBDA. Evaluation of the integer solution is based on the ambiguity success rate. Since the success rate is generally difficult to compute, some easy-to-use approximations and bounds are provided by the Ps-LAMBDA software. This success rate tool is valuable not only for inferring whether to fix the ambiguities but also for design and research purposes. For the integer estimation, the new version LAMBDA software provides more options of integer estimation and integer search, including the search-and-shrink strategy. In addition, the ratio test is incorporated to validate the significance of the fixed solution. Using these two software tools together allows for the combined execution of integer estimation and evaluation, thus benefiting multi-frequency, multi-GNSS applications.

Keywords LAMBDA · Ps-LAMBDA · Search-and-shrink · Ambiguity success rate · Ambiguity resolution

B. Li (✉) · P. J.G.Teunissen
GNSS Research Center, Department of Spatial Sciences, Curtin University, Perth, Australia
e-mail: bofeng.li@curtin.edu.au; bofeng_li@163.com

S. Verhagen
Mathematical Geodesy and Positioning, Delft University of Technology,
Delft, The Nertherlands

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_26,

26.1 Introduction

All high-precision Global Navigation Satellite Systems (GNSS) applications commonly rely on the very precise GNSS carrier-phase observations with successfully fixed ambiguities [1–5]. Hence ambiguity resolution (AR) is the key for precision GNSS applications; it comprises ambiguity estimation and evaluation.

A variety of AR methods have been developed since the late 1980s [1–6], of which the Least squares AMBiguity Decorrelation Adjustment (LAMBDA) method has become one of the more popular methods. The method includes a numerically efficient implementation of the Integer Least-Squares (ILS) principle and as such maximizes the probability of successful integer estimation [7]. The key of the LAMBDA method is to find the integer solution based on a decorrelated float ambiguity solution instead of the original one. By means of the decorrelating ambiguity transformation the efficiency of the integer search is significantly improved. For more details on the LAMBDA method, one can refer to [6, 8, 9].

High precision GNSS positioning is only possible if the integer ambiguities are correctly fixed. If the fixed solution is unreliable, it too often leads to unacceptable errors in the positioning result. Therefore, in practice, one should first evaluate (predict) how reliable the fixed solution will be. If the reliability of the integer solution is predicted to be lower than a required threshold, one should not proceed with the AR. To predict the reliability of the integer solution, the success rate is employed, i.e. the probability of correct integer estimation [7, 10, 11]. Unfortunately, exact evaluation of the success rate is generally not feasible. It is therefore necessary to find some good approximations of the success rate. So far, a variety of success rate approximations and bounds have been developed for ILS, integer bootstrapping (IB) and integer rounding (IR) [10–13]. However, up to now, no standard software has been available for conducting such evaluations.

In this contribution, we will introduce two Matlab software tools, Ps-LAMBDA and version 3.0 of LAMBDA. In the Ps-LAMBDA software, lower and upper bounds, as well as Monte-Carlo based approximations, of the ILS, IB and IR ambiguity success rates are provided. This success rate tool is valuable in applications not only for deciding on whether or not to fix the ambiguities but also for design and research purposes. In version 3.0 of LAMBDA, the following new features are incorporated: (1) more ILS search options, including a search-and-shrink; (2) extension of integer estimation methods, namely IB, IR and partial ambiguity resolution (PAR); and (3) the ratio test for validating the significance of the integer solution.

26.2 The Four Steps of Integer Ambiguity Resolution

Consider the GNSS mixed integer linear model

$$\mathbf{y} = \mathbf{A}\mathbf{a} + \mathbf{B}\mathbf{b} + \mathbf{e} \tag{26.1}$$

where $\mathbf{a}\in\mathbb{Z}^n$ and $\mathbf{b}\in\mathbb{R}^p$ are the integer and real parameter vectors, respectively; Their design matrices are $\mathbf{A}\in\mathbb{R}^{m\times n}$ and $\mathbf{B}\in\mathbb{R}^{m\times p}$ with $[\mathbf{A}\ \mathbf{B}]$ full column rank. $\mathbf{y}\in\mathbb{R}^m$ is the observation vector contaminated by the normally distributed random noise $\mathbf{e}$ with zero means and covariance matrix $\mathbf{Q}_{\mathbf{yy}}$. This mixed integer model is usually solved in four steps.

1. **Float solution**: The integer nature of the ambiguities is discarded and standard least-squares parameter estimation is performed. As a result, one obtains the so-called float solution, together with its covariance matrix

$$\begin{bmatrix}\hat{\mathbf{a}}\\ \hat{\mathbf{b}}\end{bmatrix}\sim N\left(\begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix},\begin{bmatrix}\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}} & \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{b}}}\\ \mathbf{Q}_{\hat{\mathbf{b}}\hat{\mathbf{a}}} & \mathbf{Q}_{\hat{\mathbf{b}}\hat{\mathbf{b}}}\end{bmatrix}\right) \tag{26.2}$$

2. **Integer solution**: The purpose of this step is to take the integer constraints $\mathbf{a}\in\mathbb{Z}^n$ into account. Hence, a mapping $I:\mathbb{R}^n\mapsto\mathbb{Z}^n$ is introduced that maps the float ambiguities to corresponding integer values

$$\breve{\mathbf{a}}=I(\hat{\mathbf{a}}) \tag{26.3}$$

   Many such integer mappings I exist. Popular choices are IR, IB and ILS. ILS is optimal, as it can be shown to have the largest success rate of all integer estimators. IR and IB, however, can also perform quite well, in particular after the LAMBDA decorrelation has been applied. Their advantage over ILS is that no integer search is required.

   The expected quality of $\breve{\mathbf{a}}=I(\hat{\mathbf{a}})$, as described by the ambiguity success rate, can be evaluated with the Ps-LAMBDA tool (see next section).

3. **Accept/reject**: Once integer estimates of the ambiguities have been computed, the third step consists of deciding whether or not to accept the integer solution. Several such tests have been proposed and they are all of the form

$$\text{Accept } \breve{\mathbf{a}} \text{ if } T(\breve{\mathbf{a}})<c \tag{26.4}$$

   with testing function $T:\mathbb{R}^n\mapsto\mathbb{R}$. The positive scalar $c<1$ is a tolerance value selected by the user. Thus the integer solution $\breve{\mathbf{a}}$ is accepted when $T(\breve{\mathbf{a}})$ is sufficiently small; otherwise, it is rejected in favour of the float solution $\hat{\mathbf{a}}$. Different choices for T can be made. Examples include the ratio-test, the difference-test and the projection-test [14–18]. All these tests can be cast in the framework of integer aperture estimation [19–21].

4. **Fixed solution**: In the final step, once $\breve{\mathbf{a}}$ is accepted, the float estimator $\hat{\mathbf{b}}$ is re-adjusted to obtain the so-called fixed estimator

$$\breve{\mathbf{b}}=\hat{\mathbf{b}}-\mathbf{Q}_{\hat{\mathbf{b}}\hat{\mathbf{a}}}\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}^{-1}\left(\hat{\mathbf{a}}-\breve{\mathbf{a}}\right) \tag{26.5}$$

This solution has a quality that is commensurate with the high precision of the phase data, *provided* that the uncertainty in the random integer vector $\breve{\mathbf{a}}$ can be

neglected. To decide whether or not the uncertainty in $\breve{\mathbf{a}}$ can be neglected, an evaluation of its probabilistic properties is essential.

26.3 Ps-LAMBDA Software for Ambiguity Success Rate Evaluation

26.3.1 The Ambiguity Success Rate

With the Ps-LAMBDA tool one can evaluate the expected quality of $\breve{\mathbf{a}} = I(\hat{\mathbf{a}})$, as described by its success rate. The success rate of $\breve{\mathbf{a}}$ is defined as the integral of the probability density function (PDF) of the float solution over the pull-in region $S_{\mathbf{a}}$, where $\mathbf{a}$ is the correct integer vector:

$$P_{\mathrm{s}} = P\big(\breve{\mathbf{a}} = \mathbf{a}\big) = P(\hat{\mathbf{a}} \in S_{\mathbf{a}}) = \int_{S_{\mathbf{a}}} f_{\hat{\mathbf{a}}}(\mathbf{x}|\mathbf{a})\ \mathrm{d}\mathbf{x} \tag{26.6}$$

The PDF of the float solution is

$$f_{\hat{\mathbf{a}}}(\mathbf{x}|\mathbf{a}) = \frac{1}{\sqrt{\det(2\pi\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}})}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{a})^T \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}^{-1}(\mathbf{x}-\mathbf{a})\right\} \tag{26.7}$$

As the pull-in regions of the integer estimators are integer translation invariant, the success rate can be evaluated, without knowledge of the unknown ambiguity vector, as:

$$P_{\mathrm{s}} = \int_{S_{\mathbf{0}}} f_{\hat{\mathbf{a}}}(\mathbf{x}|\mathbf{0})\,\mathrm{d}\mathbf{x} \tag{26.8}$$

This shows that the success rate depends on the PDF and on the pull-in region. The PDF is captured by covariance matrix $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$, while the pull-in region is specified by the chosen integer estimation method. For more detail on the pull-in regions of IR, IB and ILS, one can refer to [7]. For the success rate evaluation, only the covariance matrix $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$ is needed and not the float ambiguity solution $\hat{\mathbf{a}}$ itself.

26.3.2 Success Rate Approximation and Bounds

First of all, one can evaluate the success rate (26.8) by making use of Monte-Carlo simulations. The procedure is as follows. One generates a random sample $\hat{\mathbf{a}}$ from the distribution $\mathrm{N}(\mathbf{0}, \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}})$ and uses it as input for integer estimation. If the output of this estimator equals the null vector, then it is correct, otherwise it is incorrect. This process can be repeated an N number of times, and one can count how many

times the null vector is obtained as a solution, say N_s times. The approximation of the success rate follows then as:

$$P_s = N_s/N \tag{26.9}$$

In order to get good approximations, the number of samples N must be sufficiently large. The disadvantage is that it may be very time-consuming, especially in case of ILS, since for each sample an integer search is required. Therefore the Ps-LAMBDA tool also provides easy-to-compute lower and upper bounds of the ambiguity success rate.

So far there are a variety of lower and upper bounds in literature. Some of them are based on simplifying the complete covariance matrix of ambiguities and some of them based on simplifying the complicated pull-in region [22]. Besides, the relation of success rate between ILS, IB and IR methods can be used [7]:

$$P\left(\breve{\mathbf{a}}_{\mathrm{IR}} = \mathbf{a}\right) \leq P\left(\breve{\mathbf{a}}_{\mathrm{IB}} = \mathbf{a}\right) \leq P\left(\breve{\mathbf{a}}_{\mathrm{ILS}} = \mathbf{a}\right) \tag{26.10}$$

Since this ordering is the same as the ordering in terms of complexity, one may use IB success rate as lower bound of ILS and upper bound of IR or use IR as lower bound of IB. The extensive experience studies indicated that the IR and IB bounds work well based on the decorrelated ambiguities, especially, the IB success rate is a sharp lower bound of ILS. An evaluation for some of the bounds was made in [13, 23].

26.3.3 Ps-LAMBDA Software and its Demonstration

The Monte-Carlo based simulations, as well as all lower and upper bounds of the success rate are now implemented in the Ps-LAMBDA software for ILS, IB and IR methods. Figure 26.1 gives an overview of the software's structure. The main routine is **SuccessRate** which needs the inputs:

Qa	The variance matrix of the float solution $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$
method	1 = ILS [DEFAULT], 2 = IB, 3 = IR
option	The approximation/bound to compute (see Fig. 26.1)
decor	1 = decorrelation [DEFAULT], 0 = no decorrelation
nsamp	Number of samples used for simulation-based approximation

The choice for **decor** is available only to IR and IB, since these estimators are not Z-invariant. Decorrelation is always applied for ILS to ensure computational efficiency.

Figure 26.2 shows the performance of the Monte-Carlo simulation based success rate, the IB based lower bound and the region based upper bound. They are corresponding to options 1, 3 and 7 in Fig. 26.1. This result is for a CORS baseline AR with ionosphere-weighted model. The standard deviation (STD) of single-differenced ionosphere constraint is 7 cm. The 24 h dual-frequency GPS

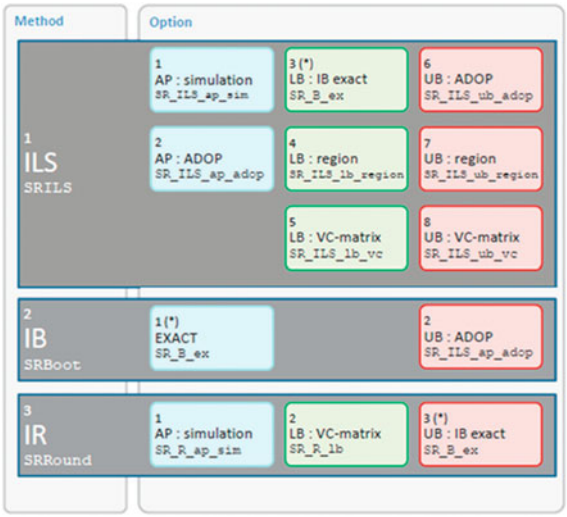


Fig. 26.1 Ps-LAMBDA: overview of available methods and options in routine **SuccessRate**. Default option is indicated with (*). Names of underlying routines are shown as well. AP = approximation, LB = lower bound, UB = upper bound

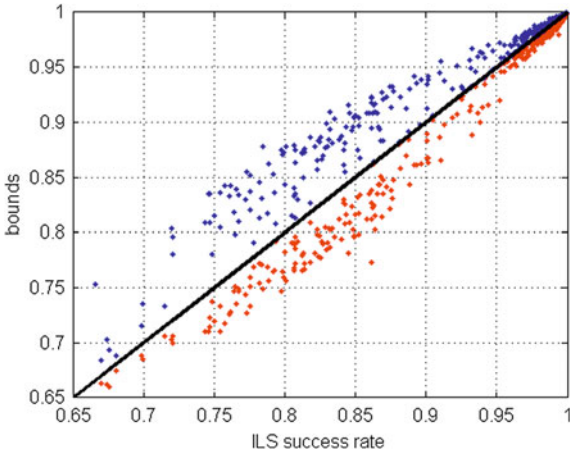


Fig. 26.2 ILS success rates: upper bound based on region (*blue*) and lower bound based on IB (*red*) versus the actual ILS success rate

observations are used with STDs of phase and code 20 cm and 2 mm, respectively. The number of epochs varies from 1 to 5 with 5 s sampling interval to obtain the success rate results from 0.65 to 1. The number of samples used for the simulation based success rate is 10^6. With this number, the approximation will be very close to the true value [24]. The result shows that both lower and upper bounds perform very well, particularly when the success rate is larger than 0.95. Considering that in the real applications, one usually accepts the fixed solution only when its success rate is close to 1 (say, >99 %), one can use these two bounds to easily evaluate the success rate. For more information about the performance of the other approximations and bounds, one can refer to [13, 23, 24].

26.4 LAMBDA (Version 3) for Ambiguity Estimation and Validation

In the previous educational version of the LAMBDA software, only the ILS method was implemented and the search was executed by enumerating all integer candidates inside the search ellipsoid with a fixed ellipsoid size. This search strategy can become time-consuming in the high-dimensional case with multi-frequency, multi-GNSS applications. Moreover, one may alternatively prefer simpler estimation method, e.g., IB or IR, in case their success rate is high enough. The main features of the LAMBDA (version 3) software are as follows:

- Additional to the enumeration with fixed ellipsoid size, another search strategy was embedded based on searching in an alternating way around the conditional estimates and shrinking the search ellipsoid. This concept was already presented in [6], see also [8], but not implemented in the previous educational versions of LAMBDA.
- As alternative to ILS, one may now also choose the IB or IR estimators.
- It is possible to output the IB success rate. This success rate is known to be a tight lower bound of the ILS success rate [7].
- PAR can be applied based on fixing a subset of the decorrelated ambiguities such that the success rate is not smaller than the required success rate [25].
- The Ratio Test can be applied to validate the significance of the fixed solution. The model-driven Fixed Failure-rate Ratio Test (FF-RT), besides the Ratio Test with a fixed (user-defined) threshold, can be applied. In FF-RT, the critical value can be automatically determined and it varies with changing model strength. For more information, one refers to [19, 21].

We now describe and geometrically illustrate the concept of the search-and-shrink approach. For the other features, one may refer to the listed references and the software manual.

26.4.1 ILS Search Space

Using the full information of $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$, the ILS AR is defined as

$$\breve{\mathbf{a}} = \arg\min (\mathbf{z} - \hat{\mathbf{a}})^T \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}^{-1} (\mathbf{z} - \hat{\mathbf{a}}),\ \forall \mathbf{z} \in \mathbb{Z}^m \tag{26.11}$$

The integer minimizer (26.11) is obtained through a search over the integer grid points of an n-dimensional hyper-ellipsoid defined by

$$F(\mathbf{z}) = (\mathbf{z} - \hat{\mathbf{a}})^T \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}^{-1} (\mathbf{z} - \hat{\mathbf{a}}) \leq \chi^2 \tag{26.12}$$

The integer grid point $\mathbf{z}$ inside the hyper-ellipsoid which gives the minimum value of function $F(\mathbf{z})$ is the optimal ILS solution $\breve{\mathbf{a}}$. The search efficiency is

governed by the size χ^2 and the shape of ellipsoid. The constant χ^2 can be predetermined using different strategies and can also be shrunken during the search [8, 26]. The shape and orientation of the ellipsoid are defined by the covariance matrix $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$. The high correlated covariance matrix will often lead to search halting. To improve the search efficiency, the decorrelated variables, $\hat{\mathbf{z}} = \mathbf{Z}^T\hat{\mathbf{a}}$ and $\mathbf{Q}_{\hat{\mathbf{z}}\hat{\mathbf{z}}} = \mathbf{Z}^T\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}\mathbf{Z}$ are used in (26.12) instead of $\hat{\mathbf{a}}$ and $\mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}}$, respectively.

26.4.2 Search-and-Shrink in Two Dimensions

In order to explain the concept of the search-and-shrink procedure, we describe it with a two-dimensional (2D) example. Figures 26.3, 26.4 show how the search procedures work with enumeration and shrinking for a 2D example. The float solution is depicted with a *blue* asterisk. The grey line shows the line defined as

$$\hat{z}_{1|2} = \hat{z}_1 - \sigma_{\hat{z}_1\hat{z}_2}\sigma^{-2}_{\hat{z}_2\hat{z}_2}(\hat{z}_2 - z_2) \tag{26.13}$$

It shows that if $\hat{z}_2$ is rounded to its nearest integer 0, the conditional estimate $\hat{z}_{1|2}$ will be the intersection of the grey line with the grid line at $z_2 = 0$. With bootstrapping, this conditional estimate is then rounded to its nearest integer 1. If only 2 integer vectors are requested, both enumeration and shrinking search strategies will give the same initial search ellipse, see left panel of Fig. 26.3. If 6 candidate integer vectors are requested, the difference between the two search strategies becomes clear. Figure 26.3 shows on the right the search ellipse for the enumerating search strategy. All 16 grid points inside ellipse are examined to choose 6 candidates (*black dots*) with the smallest $F(\mathbf{z}^{(i)})$.

With the search-and-shrink strategy, the 6-step procedure is described and each step is shown as a separate panel in Fig. 26.4. For each step the new candidate is

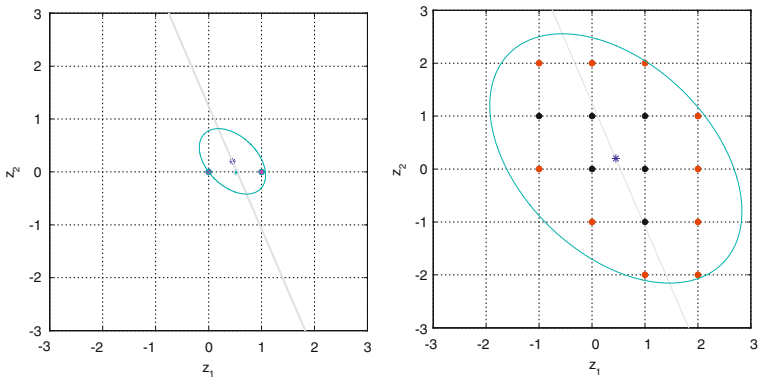


Fig. 26.3 Search ellipse for 2D example for enumerating search strategy. The requested number of integer vectors is 2 (*left*) or 6 (*right*)

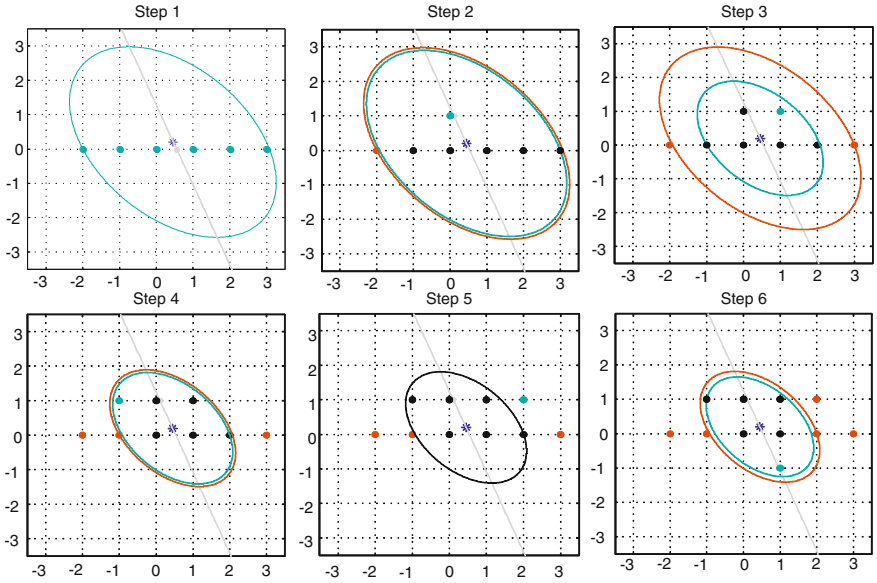


Fig. 26.4 Search-and-shrink procedure for 2D example (The x and y axes denote the ambiguity z_1 and z_2, respectively). The requested number of integer vectors is 6

shown in *blue* in the corresponding panel. Candidates from previous steps are shown in *black*. If at a certain stage a candidate is removed, as it is outside the shrunken ellipse, it is shown in *red*. For every step the old search ellipse (*red*) and the new shrunken ellipse (*blue*) are shown if shrinking is possible.

1. With the bootstrapped solution $[1\ 0]^T$ as the first candidate, we find the next 5 candidates by rounding the conditional ambiguity $\hat{z}_{1|2}$ to its 2nd, 3rd, 4th and 5th nearest integer. We now have 6 candidates $\mathbf{z}(i)$ (blue points), and the size of the search ellipse is set to the maximum $F(\mathbf{z}(i))$, see the blue ellipse.
2. Round $\hat{z}_2$ to its second nearest integer ($z_2 = 1$), and round the corresponding new conditional estimate of the first ambiguity to the nearest integer. This will give you a new candidate (blue point in panel 2). It resides in the search ellipse. Therefore, the integer candidate with the largest $F(\mathbf{z}(i))$ is removed (see red point), and the ellipse size is now set to the largest value of the remaining candidates.
3. Next, round the conditional estimate of the first ambiguity from Step 2 to the second nearest integer. Again it resides in the shrunken search ellipsoid. Therefore, the integer candidate with the largest $F(\mathbf{z}(i))$ is removed, and the ellipse size is now set to the largest value of the remaining candidates.
4. Round the conditional estimate of the first ambiguity from Step 2 to the third nearest integer. Again it resides in the shrunken search ellipse. Therefore, the integer candidate with the largest $F(\mathbf{z}(i))$ is removed, and the ellipse size is now set to the largest value of the remaining candidates.

5. Round the conditional estimate of the first ambiguity from Step 2 to the fourth nearest integer. This candidate is outside the search ellipsoid and is disregarded. No shrinking in this step.
6. A new candidate for the second ambiguity is obtained by rounding $\hat{z}_2$ to the third nearest integer ($z_2 = -1$). Rounding the corresponding new conditional estimate of the first ambiguity to the nearest integer results in a new candidate, this resides in the search ellipse. Therefore, the integer candidate with the largest $F(\mathbf{z}(i))$ is removed, and the ellipse size is now set to the largest value of the remaining candidates.

After this step, no further shrinking is possible, and the requested 6 candidates have been obtained.

For the n-dimensional case the search-and-shrink strategy goes along similar lines.

26.5 Summary

In this paper we introduced two Matlab software tools, Ps-LAMBDA and LAMBDA (version 3). They are developed for integer ambiguity evaluation and estimation, respectively. With Ps-LAMBDA the ambiguity success rates, and their bounds, of ILS, IB and IR can be computed, while the new version of LAMBDA provides more options for integer estimation, including the fixed failure-rate ratio test. The Ps-LAMBDA and new version LAMBDA softwares can be downloaded from http://gnss.curtin.edu.au/.

Acknowledgments This work was done in the framework of the project 1.01 'New Carrier-Phase Processing Strategies for Next Generation GNSS Positioning' of the Cooperative Research Centre for Spatial Information. Professor Teunissen is the recipient of an Australian Research Council Federation Fellowship (project No. FF0883188). This support is gratefully acknowledged.

References

1. Hofmann-Wellenhof B, Lichtenegger H, Collins J (2001) Global positioning system: theory and practice, 5th edn. Springer Berlin Heidelberg, New York
2. Leick A (2004) GPS satellite surveying, 3rd edn. John Wiley, New York
3. Teunissen PJG, Kleusberg A (1998) GPS for geodesy, 2nd edn. Springer, Berlin Heidelberg New York
4. Blewitt G (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. J Geophys Res 94:135–151
5. Dong D, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. J Geophys Res 94 (B4): 3949–3966

6. Teunissen PJG (1993) Least squares estimation of the integer GPS ambiguities. In: Invited lecture, section IV theory and methodology, IAG General Meeting, Beijing
7. Teunissen PJG (1999) An optimality property of the integer least-squares estimator. J Geodesy 73(11):587–593
8. De Jonge PJ, Tiberius CCJM (1996) The LAMBDA method for integer ambiguity estimation: implementation aspects, LGR-Series, No 12. Tech. report, Delft University of Technology
9. Li B, Teunissen PJG (2011) High dimensional integer ambiguity resolution: A first comparison between LAMBDA and Bernese. J Navig 64:S192–S210
10. Hassibi A, Boyd S (1998) Integer parameter estimation in linear models with applications to GPS. IEEE Trans Sig Process 46(11):2938–2952
11. Teunissen PJG (1998) Success probability of integer GPS ambiguity rounding and bootstrapping. J Geodesy 72:606–612
12. Teunissen PJG (2000) ADOP based upper bounds for the bootstrapped and the least-squares ambiguity success rates. Artif Satell 35(4):171–179
13. Verhagen S (2005) On the reliability of integer ambiguity resolution. Navigation 52(2):99–110
14. Euler HJ, Schaffrin B (1991) On a measure for the discernibility between different ambiguity solutions in the static-kinematic GPS-mode. In: Proceedings of Kinematic Systems in Geodesy, Surveying, and Remote Sensing, International Association of Geodesy Series, vol 107. pp 285–295
15. Han S (1997) Quality control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. J Geodesy 71(6):351–361
16. Landau H, Euler HJ (1992) On-the-fly ambiguity resolution for precise differential positioning. In: Proceedings of ION GPS-1992, Albuquerque NM, pp 607–613
17. Tiberius CCJM, De Jonge P (1995) Fast positioning using the LAMBDA method. In Proceedings of DSNS'95, Bergen, Norway, paper no. 30. The Nordic Institute of Navigation, Oslo
18. Wang J, Stewart MP, Tsakiri M (1998) A discrimination test procedure for ambiguity resolution on-the-fly. J Geodesy 72(11):644–653
19. Teunissen PJG, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. Surv Rev 41(312):138–151
20. Verhagen S, Teunissen PJG (2006) New global navigation satellite system ambiguity resolution method compared to existing approaches. J Guidance Control Dyn 29(4):981–991
21. Verhagen S, Teunissen PJG (2012) The ratio test for future GNSS ambiguity resolution. GPS Solut. doi:10.1007/s10291-012-0299-z
22. Teunissen PJG (1998) On the integer normal distribution of the GPS ambiguities. Artif Satell 33(2):49–64
23. Thomsen HF (2000) Evaluation of upper and lower bounds on the success probability. In: Proceedings of ION GPS-2000, Salt Lake City UT, pp 183–188
24. Verhagen S, Li B, Teunissen PJG (2013) Ps-LAMBDA: ambiguity success rate evaluation software for interferometric applications. Comput Geosci 54:361–376
25. Verhagen S, Teunissen PJG, van der Marel H, Li B (2011) GNSS ambiguity resolution: which subset to fix? IGNSS Symposium 2011, Sydney, Australia
26. Chang X, Yang X, Zhou T (2005) MLAMBDA: a modified LAMBDA method for integer least-squares estimation. J Geodesy 79:552–565

Chapter 27
The Method to Inverse PWV Using VMF1 Grid Data

Min Wong, Hongzhou Chai, Zongpeng Pan and Yanli Chen

Abstract In this paper, the method interpolating Vienna Mapping Function (VMF1) grid data to estimate Precipitable Water Vapor (PWV) is introduced. Using precise point positioning technique, the PWV values are estimated without external meteorological measurements. The experiment results demonstrate a good agreement with those derived with current state-of-the-art procedure that relies on site-specific meteorological observations. The zenith path delay (ZPD) results are compared with International GNSS Service (IGS) troposphere product and the RMS of differences are within 8.3 mm. Compared with sounding data, the RMS of the PWV difference are less than 5 mm, which preliminarily validates the feasibility to apply this method to regional weather monitoring.

Keywords Precipitable water vapor · Vienna Mapping Function 1 (VMF1) · Precise point positioning

27.1 Introduction

Since the technique to inverse the Precipitable Water Vapor (PWV) from Global Positioning System (GPS) signal was come up two decades ago [2], the method has been widely applied because of its advantages such as high spatial and temporal resolution, low observation cost and immunity from bad weather condition. However, the conventional ground GPS-based PWV estimation needs ground

M. Wong (✉) · H. Chai · Z. Pan
Institute of Geospatial Information, Information Engineering University, Zhengzhou 450052, China
e-mail: difzibo@163.com; different9@163.com

Y. Chen
68011 Troops of PLA, Lanzhou 730030, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_27,

meteorological parameter derived with external sensors, which increases the construction cost of observation station and limits its application range.

The applications of Numerical Weather Models (NWM) provided by the United States National Centers for Environmental Prediction (NCEP) and the European Center for Medium-Range Weather Forecasts (ECMWF) have provided an alternative way to obtain surface meteorological data. Schüler [8] validate the possibility to estimate PWV using meteorological parameters interpolated from NCEP Global Data Assimilation System (GDAS) reanalysis data. Jade and Vijayan [7] compute the PWV time series in Indian subcontinent based on NCEP reanalysis data, and the accuracy of PWV values is comparable to those from radiosonde measurement. Chang and He [5] perform a similar technique to estimate high accuracy PWV values over Shanghai, which shows the technique can be helpful to the precipitation forecast of the hydropower station with GPS even when nearby meteorological sensors are unavailable.

The Vienna Mapping Function (VMF1) grid data are released by the University of Technology of Vienna, which contain the Zenith Hydrostatic Delay (ZHD) and Zenith Wet Delay (ZWD) derived from ECMWF-NWM and parameters to calculate VMF1 on a $2° \times 2.5°$ global grid. The gird data have facilitated some high accuracy geodetic applications over past years. And they provide an easy access to use ECMWF-NWM information without complicated computation. However, the possibility to use the grid data for PWV estimation has been reported in little work. In this paper, a method to estimate PWV values without meteorological measurements is presented. The method spatially interpolates the VMF1 grid data and then processes the GPS data using precise point positioning (PPP) technique. By comparing with external references and results of conventional method that relies on surface meteorological sensors, the accuracy of estimated troposphere zenith path delay (ZPD) and PWV values is assessed.

27.2 GPS PWV Inversion

The ZPD along GPS signal propagation can be divided as Eq. (27.1):

$$ZPD = ZHD \cdot M_{\mathrm{h}}(\theta) + ZWD \cdot M_{\mathrm{w}}(\theta) + (Gn \cdot \cos(\alpha) + Ge \cdot \sin(\alpha)) \cdot M_{\mathrm{g}}(\theta) \tag{27.1}$$

Where $M_{\mathrm{h}}(\theta)$ is the hydrostatic mapping function, $M_{\mathrm{w}}(\theta)$ is the wet mapping function, and $M_{\mathrm{g}}(\theta)$ is the gradient mapping function. Gn, Ge are the horizontal delay gradient components. θ is the satellite elevation angle and α is satellite azimuth angle. Within the PPP computation process, ZHD is stable and easy to model with empirical troposphere refraction model, whereas, ZWD is estimated as random walk process according to its stochastic characteristic. Using below linear relationship, the estimated ZWD can be converted to PWV [1]:

$$PWV = \prod \cdot ZWD \tag{27.2}$$

where $\prod$ is the conversion coefficient computed with following Eq. (27.3):

$$\prod = 10^6 [R_v \cdot \rho_l \cdot (\frac{k_3}{T_m} + k_2')]^{-1} \tag{27.3}$$

where R_v is the gas constant for the water vapor, ρ_l is the density of liquid water, k_2', k_3 are physical constant, T_m is the weighted mean temperature of the atmosphere, which can be derived with the linear equation related to station surface temperature. In this paper, the linear regression come up by Bevis el al. is used [2]:

$$T_m = 70.2(K) + 0.72 \cdot T_s(K) \tag{27.4}$$

where T_s is the station surface temperature in Kelvin.

27.3 Method to Interpolate VMF1 Grid Data

The VMF1 grid files are downloaded from University of Technology of Vienna's Website (http://mars.hg.tuwien.ac.at/~ecmwf1/ accessed on 30 October 2012). The grid files that are generated from ECMWF-NWM are based on global $2^\circ \times 2.5^\circ$ grid, and updated every 6 h. At every grid point of the grid data, there are four coefficients: wet mapping function coefficient a_w, hydrostatic mapping function coefficient a_h, ZHD z_h and ZWD prior-value z_w. The gird point height is defined in the file orography_ell (available at the same website). The a_w and a_h are fitted with the four times a day ray-tracing with the ECMWF-NWM [3], meanwhile, the function form is same as Niell Mapping Function (NMF), and is described in Eq. (27.5):

$$mf(\theta, a, b, c) = \frac{1 + \frac{a}{1+\frac{b}{1+c}}}{\sin\theta + \frac{a}{\sin\theta + \frac{b}{\sin\theta + c}}} \tag{27.5}$$

where θ is satellite elevation angle, b and c are function coefficients. User could perform the bi-linear interpolation method to obtain the four parameters at specific orography correspond to the height that we referred to as interpolation height. Normally, the interpolation height is not equal to the user site height, hence, we need to correct the z_h for the height discrepancy with the method proposed by Steigenberger et al. [9]:

$$ZHD(h_s) = ZHD(h_g) - 0.002277 \frac{g \cdot P(h_g) \cdot (h_s - h_g)}{R \cdot T(h_g)} \tag{27.6}$$

where $ZHD(h_g)$ and $ZHD(h_s)$ are ZHDs refereed to site location and interpolation point respectively, h_g is the grid point height, h_s is the site height, and both h_g and h_s are in kilometers. g is the gravity constant (9.784 m/s^2), R is the gaz constant

(0.289644 kJ/kg). $P(h_g)$ and $T(h_g)$ are pressure and temperature of grid point, here we calculate them with GPT model [6].

27.4 Results and Analysis

Observations of seven International GNSS Service (IGS) stations distributed in east Asia region (as shown in Fig. 27.1) during July 15–23, 2012(the days of year (DOY) from 2012 197 to 205) are selected in our experiment data set. The ZPD are estimated with two scenarios and converted to PWV, respectively. By comparing the results of two scenarios, the accuracy of PWV values is assessed. Moreover, the IGS troposphere product and sounding data are used as external references for ZPD and PWV results separately. The two scenarios are hereafter referred to as ECMWF scenario and SAAS scenario. The ECMWF scenario uses the ZHD, prior ZWD and VMF1 values interpolated from VMF1 grid data. The SAAS scenario is commonly practiced for GPS-based meteorological inversion, which uses the ZHD values computed with Saastamoinen model fed by surface meteorological parameters, and NMF as mapping function. Within data processing, the precise point positioning (PPP) technique is applied, the cutoff elevation is 10° and the estimation interval is 300 s. For the unavailability of IGS troposphere product of days from July 16 to 17 and July 21 for BJFS station and July 16 for DAEJ station (accessed on 15 September 2012), the estimation results of corresponding dates are not applied to comparison with external references.

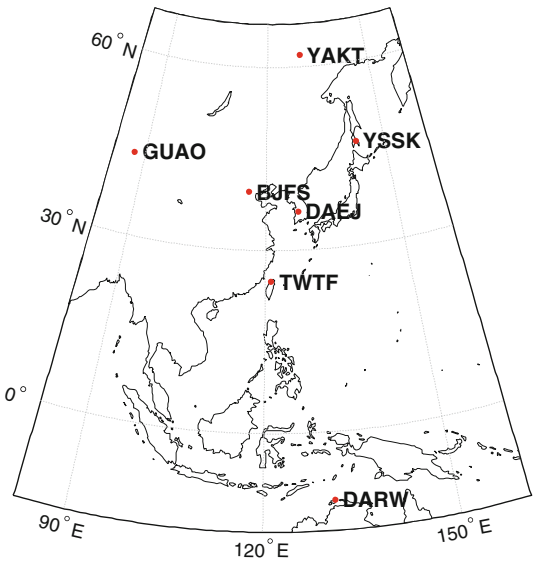


Fig. 27.1 Stations location

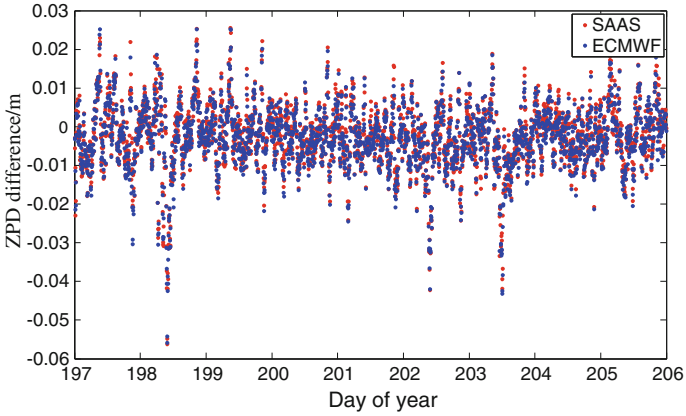


Fig. 27.2 Differences of ZPD values computed with different models and IGS troposphere product for site TWTF

27.4.1 Comparison of ZPD Results

As presented in Fig. 27.2, the ZPD results estimated with two scenarios are close to each other and both show a good agreement with IGS troposphere product. Mostly, the discrepancies from two scenarios compared to IGS troposphere product are within 10 mm, which is similar to the coordinate bias when the two scenarios are applied to position estimation [3].

The differences statistics of ZPD estimation results and IGS troposphere product for all seven stations are shown in Table 27.1. Both results from two scenarios show a good agreement with IGS troposphere product, the mean biases are less than 6 mm, and the RMS are about 7 mm, which is comparable to the accuracy range of IGS troposphere product [4]. Meanwhile, the consistency of ZPD results from two scenarios indicates they are close to each other at accuracy.

Table 27.1 Mean biases and bias RMS of ZPD differences from different models compared to IGS troposphere product

Station	SAAS		ECMWF	
	Mean bias (m)	Bias RMS (m)	Mean bias (m)	Bias RMS (m)
BJFS	−0.0057	0.0077	−0.0061	0.0076
DAEJ	−0.0050	0.0081	−0.0052	0.0081
DARW	−0.0049	0.0080	−0.0041	0.0080
GUAO	−0.0050	0.0079	−0.0047	0.0078
TWTF	−0.0024	0.0083	−0.0030	0.0083
YAKT	−0.0024	0.0068	−0.0034	0.0067
YSSK	−0.0043	0.0070	−0.0041	0.0071
Mean	−0.0042	0.0077	−0.0044	0.0077

27.4.2 Comparison of PWV Results

We then extend our study to investigate the influence of different estimation scenarios on PWV estimation and assess the accuracy of PWV results derived with VMF1 grid data. The PWV results from four different stations (BJFS, DARW, GUAO and YSSK) are compared with sounding data collected at the same city. For unavailability of corresponding sounding data for other three station used in last section, the PWV results of those are not used here. The sounding data are accessed from the University of Wyoming Department of Atmospheric Science's website (http://weather.uwyo.edu/upperair accessed on 30 October 2012). The sample rate of sounding data is two times one day, and all the sounding data sensors are not collocated with GPS station, although they are located in same city. Because of the difference of sample rate, the PWV results are averaged with a 30 min window centered at the collection points of sounding data and then applied to comparison.

The comparisons of PWV results and sounding data are presented in Fig. 27.3 and Fig. 27.4 for station BJFS and GUAO, respectively. There are minor differences of two scenarios as we can see, and the maximum difference is only 3.8 mm. Compared to sounding data, at most time, the differences of ECMWF scenario results and sounding data are less than 10 mm. Moreover, benefited from their high temporal resolution, the GPS-based results can reveal more detail information of PWV variations.

Figure 27.3 describes the PWV variation of BJFS station before and after the rainfall on July 21, 2012 (DOY 203), which is the heaviest rainfall in 61 years fell on the Chinese capital city of Beijing. The rain started in the afternoon and continued through the evening. BJFS station locates at the city's Fangshan District

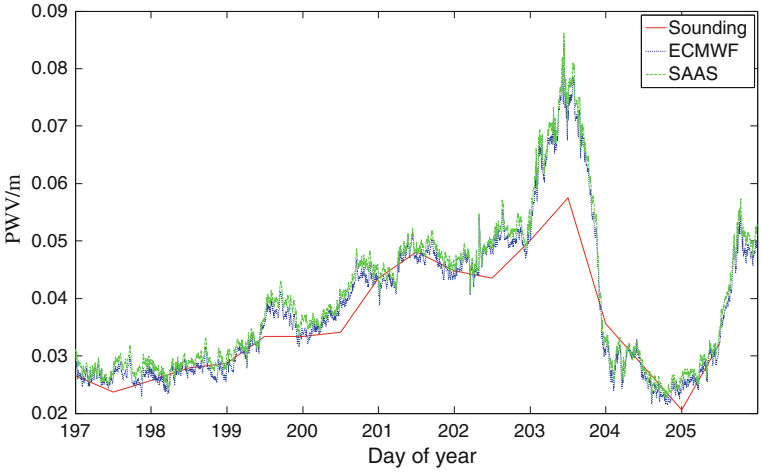


Fig. 27.3 Comparison of PWV values computed with different methods for site BJFS

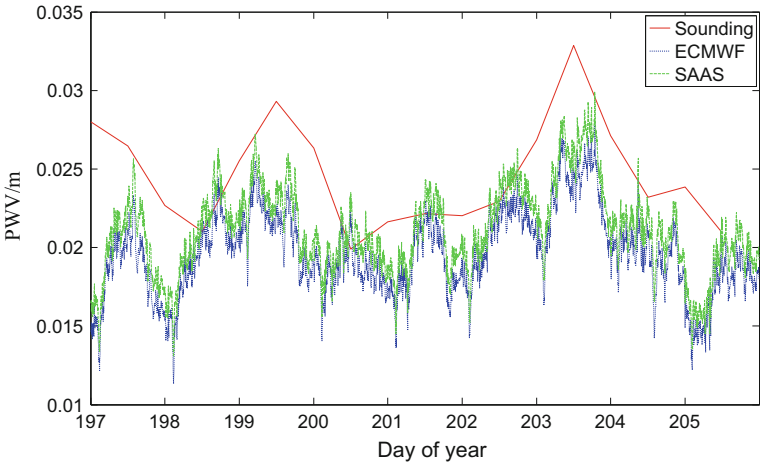


Fig. 27.4 Comparison of PWV values computed with different methods for site GUAO

that experienced the most serious rainfall. The estimated PWV values continuously increased from DOY 197, suddenly soared on DOY 203, and reached its maximum value at 10:45 UTC. And then, along with the precipitation process, the PWV values began to sharply decrease and reached its minimum value at 18:55 UTC. It is noticeable that the maximum difference of PPP results and sounding data appeared on DOY 203 when the rainfall happened. The extreme weather may contribute to this rather large difference. To the contrary, the steadier weather condition at station GUAO makes the differences of station GUAO more stable.

The statistics of differences of PWV results estimated from two scenarios are shown in Table 27.2. Even under extreme weather condition, the results from different scenarios have a good agreement that both the mean bias and RMS are less than 5 mm. The millimeter level differences of ECMWF scenario and SAAS scenario validate the feasibility of ECMWF scenario to derive high accuracy PWV results without ground surface meteorological data.

Table 27.2 Mean biases and bias RMS of differences between GPS-derived PWV results and sounding data

Station	ECMWF		SAAS	
	Mean bias (m)	Bias RMS (m)	Mean bias (m)	Bias RMS (m)
BJFS	0.0016	0.0039	0.0033	0.0040
DARW	−0.0019	0.0028	−0.0019	0.0027
GUAO	−0.0048	0.0038	−0.0033	0.0037
YSSK	0.0018	0.0045	0.0022	0.0043

27.5 Conclusions

In this paper, a method to apply VMF1 grid data to PWV inversion without ground surface meteorological data is presented. The experiment results demonstrated their comparable accuracy to the conventional method relies on external meteorological sensors. Comparing to sounding data, the RMS of difference is within 8.3 mm, which indicates the high accuracy of PWV values estimated with VMF1 grid data. This method reveals its application potential to decrease construction cost of GPS meteorology observation station and then to exert its advantage in high spatial and temporal resolution.

The future study will focus on the investigation of large difference conincided with heavy rainfall and adoption of forecast VMF1 grid data to improve the capability of real-time weather monitoring.

Acknowledgments This work is supported by China National Natural Science Foundation of China (No: 41274045).

References

1. Bevis M, Businger S, Chiswell S et al (1994) GPS meteorology: mapping zenith wet delays onto precipitable water. J Appl Meteorol 33(3):379–386
2. Bevis M, Businger S, Herring T et al (1992) GPS meteorology—remote sensing of atmospheric water vapor using the Global Positioning System. J Geophys Res 97 (D14):15787–15801
3. Boehm J, Werl B, Schuh H (2006) Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. J Geophys Res 111 (B2):B02406
4. Byun SH, Bar-Sever YE (2009) A new type of troposphere zenith path delay product of the international GNSS service. J Geodesy 83(3):1–7
5. Chang L, He X (2010) Regional precipitation forecast using GPS and NCEP. Scientia Sinica Phys, Meth Astron 40(5):685–692 (in Chinese)
6. Fund F, Morel L, Mocquet A (2009) Discussion and recommendations about the height correction for a priori zenith hydrostatic delays derived from ECMWF data
7. Jade S, Vijayan MSM (2008) GPS-based atmospheric precipitable water vapor estimation using meteorological parameters interpolated from NCEP global reanalysis data. J Geophys Res 113 (D3):D03106
8. Schüler T (2001) On ground-based GPS tropospheric delay estimation. Ph.D. Thesis, Universtät der Bundeswehr Müchen, Universitätsbibliothek, Neubiberg, Germany
9. Steigenberger P, Boehm J, Tesmer V (2009) Comparison of GMF/GPT with VMF1/ECMWF and implications for atmospheric loading. J Geodesy 83(10):943–951

Chapter 28
The Analysis of Ill Posedness in GNSS High-Precision Differential Positioning

Wang Gao, Chengfa Gao, Shuguo Pan, Shengli Wang and Denghui Wang

Abstract GNSS high-precision differential positioning generally requires the high-precision float solution and the reasonable covariance matrix for ambiguity integers in short observational time spans. But station-star geometric changes little in short observational time spans, which causes the design matrix to be ill-conditioned seriously during adjustment calculation and leads to the solving of parameters is unstable. With the development of Compass Navigation Satellite System, the GNSS high-precision positioning is now entering the multi-system combined ages. GPS, GLONASS and Compass have different orbits and operating features so that there is some difference from each other in adjustment calculation. The paper combining with the analysis theory on ill-conditioned system, analyses the ill posedness of GPS, GLONASS and Compass comparatively through a method called spectrum analysis of eigenvalue. The paper reveals the characteristics of the ill condition in several different positioning forms. The conclusion that the faster the satellites runs, the weaker the ill posedness of the parameters solving system will be is presented.

W. Gao (✉) · C. Gao · D. Wang
School of Transportation, Southeast University, 2 Sipailou 210096 Nanjing, China
e-mail: gaowang_1990@126.com; 1300317284@qq.com

C. Gao
e-mail: gaochfa@163.com

D. Wang
e-mail: owenxun@gmail.com

S. Pan · S. Wang
School of Instrument Science and Engineering, Southeast University, 2 Sipailou 210096 Nanjing, China
e-mail: panshuguo@gmail.com

S. Wang
e-mail: victory_wsl@126.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_28,

Keywords GNSS double-differential positioning · Ill posedness · Algebraic structure · Spectrum analysis

28.1 Introduction

With the establishment and being putting into use gradually of Compass Navigation Satellite System, the GNSS high-precision positioning is now entering the multi-system combined ages [1]. There are certain differences in satellite constellations of GPS, GLONASS and Compass [2, 3]. GPS has six satellite orbital planes and each orbit has an average height of about 20,200 km. The satellite operation cycle is about 11 h and 58 min. GLONASS has 3 satellite orbital planes and each orbit has an average height of 19,100 km. The satellite operation cycle is about 11 h 15 min. Compass regional navigation system is different from GPS, GLONASS and other GNSS system because its constellation design not only includes the MEO satellites but also includes 5 GEO satellites and 5 IGSO satellites. GEO and IGSO satellites have run in the network and 4 MEO satellites have already been launched step by step. GEOs are geostationary satellites and IGSOs are inclined geostationary observation satellites. Both of them are high-orbit satellites and their orbital altitude is 36,000 km approximately. GNSS carrier phase high-precision positioning model is generally based on the theory of integer least squares. The parameters are usually estimated by the following three steps [4]: the float solution of parameters, the search of integer ambiguity and the fixed solution of ambiguity. Due to the existence of ambiguity parameters, the solution model are usually ill-posed. The ill posedness in GNSS was discussed in [5–8] and in [4] the condition-number method was adopted to describe the ill posedness quantitatively. The general method to solving the problem is to make the geometric relationship between satellites and stations change by long-term and multi-epoch observations. This change depends mainly on the variation of the angles between the stations and satellites. The rate of the change depends on the height of the satellites' orbit and operating characteristics. The difference of satellites' orbital altitude and operating characteristics between the GPS, GLONASS, and Compass cause various degrees of ill-posedness in high-precision positioning process.

In this paper, based on the theory on ill-conditioned system, a method called analysis of characteristic spectral of a matrix is represented. The method is used to analyze comparatively the characteristics and changes of the characteristic spectral of the normal matrix in the calculating process of different systems. This is divided into two modes–the mode of conventional baseline solution and the mode of network RTK solution so that the method can reveal the characteristics of algebraic structure in different systems and different patterns.

28.2 The Model of GNSS Carrier High-Precision Double-Differential Positioning

The model of DD-carrier phase observables can be written as [9, 10]:

$$\lambda\Delta\nabla\varphi_{A,B}^{p,q}(t_i) = \Delta\nabla R_{A,B}^{p,q}(t_i) + \lambda\Delta\nabla N_{A,B}^{p,q} + \delta_{A,B}^{p,q}(t_i) \tag{28.1}$$

$$\delta_{A,B}^{p,q}(t_i) = -\Delta\nabla I_{A,B}^{p,q}(t_i) + \Delta\nabla T_{A,B}^{p,q}(t_i) + \Delta\nabla\delta m_{A,B}^{p,q}(t_i) + \Delta\nabla\varepsilon_{A,B}^{p,q}(t_i) \tag{28.2}$$

Where $\Delta\nabla X_{A,B}^{p,q}(t_i)$ is double-difference value of X from receiver A, B and satellite p, q at the time t_i; λ is nominal carrier phase wavelength and φ is carrier phase observables; R is geometric range from a receiver and a satellite; N is integer ambiguity; I and T are dispersive atmospheric (ionospheric) delay and non-dispersive atmospheric delay; m is multipath on carrier; ε is noise on carrier and other errors.

Generally, there are two modes of DD-carrier phase positioning, which are the mode of baseline solution and the mode of network RTK. The error equation of baseline solution can be written as [9, 10]:

$$\begin{pmatrix} v^{1,q} \\ v^{2,q} \\ \vdots \\ v^{p,q} \end{pmatrix} = \begin{pmatrix} l^{1,q} & m^{1,q} & n^{1,q} \\ l^{2,q} & m^{2,q} & n^{2,q} \\ \vdots & \vdots & \vdots \\ l^{p,q} & m^{p,q} & n^{p,q} \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} + \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{pmatrix} \begin{pmatrix} \Delta^{1,q} \\ \Delta\nabla N^{2,q} \\ \vdots \\ \Delta\nabla N^{p,q} \end{pmatrix} - \begin{pmatrix} \Delta\nabla L^{1,q} \\ \Delta\nabla L^{2,q} \\ \vdots \\ \Delta\nabla L^{n,q} \end{pmatrix} \tag{28.3}$$

where p is the number of a satellite and q is the number of the reference satellite; L is constant term including observables; dx, dy, dz are coordinate increments from station A to station B; $l^{p,q}, m^{p,q}, n^{p,q}$ are coefficients of coordinate increment, which are usually expressed as [2]:

$$\begin{aligned} l^{p,q} &= -\frac{1}{2}\left(\frac{\Delta X_A^p}{\rho_A^p} + \frac{\Delta X_B^p}{\rho_B^p} - \frac{\Delta X_A^q}{\rho_A^q} - \frac{\Delta X_B^q}{\rho_B^q}\right) \\ m^{p,q} &= -\frac{1}{2}\left(\frac{\Delta Y_A^p}{\rho_A^p} + \frac{\Delta Y_B^p}{\rho_B^p} - \frac{\Delta Y_A^q}{\rho_A^q} - \frac{\Delta Y_B^q}{\rho_B^q}\right) \\ n^{p,q} &= -\frac{1}{2}\left(\frac{\Delta Z_A^p}{\rho_A^p} + \frac{\Delta Z_B^p}{\rho_B^p} - \frac{\Delta Z_A^q}{\rho_A^q} - \frac{\Delta Z_B^q}{\rho_B^q}\right) \end{aligned} \tag{28.4}$$

ΔX_i^n, ΔY_i^n, ΔZ_i^n are the difference of corresponding coordinate from satellite n and observation station i. The error equation of network RTK can be written as [11]:

$$\begin{pmatrix} v^{1,q} \\ v^{2,q} \\ \vdots \\ v^{p,q} \end{pmatrix} = \begin{pmatrix} \nabla MF(\theta_A^{1,q}) & -\nabla MF(\theta_B^{1,q}) \\ \nabla MF(\theta_A^{2,q}) & -\nabla MF(\theta_B^{2,q}) \\ \vdots & \vdots \\ \nabla MF(\theta_A^{p,q}) & -\nabla MF(\theta_B^{p,q}) \end{pmatrix} \begin{pmatrix} ZTD_{A,wet} \\ ZTD_{B,wet} \end{pmatrix} + \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{pmatrix} \begin{pmatrix} \Delta\nabla N^{1,q} \\ \Delta\nabla N^{2,q} \\ \vdots \\ \Delta\nabla N^{p,q} \end{pmatrix} - \begin{pmatrix} \Delta\nabla L^{1,q} \\ \Delta\nabla L^{2,q} \\ \vdots \\ \Delta\nabla L^{p,q} \end{pmatrix} \tag{28.5}$$

where $\nabla MF(\theta_i^{p,q})$ is the difference of mapping coefficients of troposphere delay (generally meaning wet delay) from satellite p and satellite q in observation station i, in this paper we adopt Niell model [12]; $ZTD_{i,wet}$ is zenith wet delay of troposphere in observation station i.

Equations (28.3) and (28.5) can be simplistically and uniformly expressed as:

$$V = BX - L \tag{28.6}$$

The least-square solution of Eq. (28.6) is

$$X = (B^T PB)^{-1} B^T PL \tag{28.7}$$

This in baseline solution can be expressed as:

$$X = \begin{bmatrix} dx & dy & dz & \Delta\nabla N^{1,k} & \Delta\nabla N^{2,k} & \cdots & \Delta\nabla N^{n,k} \end{bmatrix}^T.$$

And in network RTK solution can be expressed as:

$$X = \begin{bmatrix} \Delta\nabla N^{1,k} & \Delta\nabla N^{2,k} & \cdots & \Delta\nabla N^{n,k} & ZTD_{p,wet} & ZTD_{q,wet} \end{bmatrix}^T.$$

28.3 The Quantitative Description and Diagnostic Methods of the Ill Condition in Parameter-Estimation System

If a parameter-estimation system has serious ill posedness, the solving of parameters will be very susceptible to the influence of various factors and become unstable. As we all know the noise of observations and other errors cannot be avoided, therefore this factor will affect the efficiency and accuracy of the parameters for solving in particular for short baseline in network RTK. Reasonable measure or diagnosis for the ill condition in a parameter-estimation system is the premise of weakening or overcoming the impact of the ill condition. There are several common methods for

measuring or diagnosing ill condition [13]: the method of condition number, the method of eigenvalue analysis, the method of variance inflation factor (VIF) and the method of determinant discriminant. In GNSS high-precision double-differential carrier positioning, the equation will no longer rank-defect after several-epoch observations generally. At the time, the normal matrix N is symmetric and positive definite and the condition number is the ratio of the maximum and minimum eigenvalues, namely:

$$cond(N) = \frac{\lambda_{max}}{\lambda_{min}} \tag{28.8}$$

where λ_{max} and λ_{min} are the maximum and the minimum eigenvalue respectively.

The condition-number method mainly focuses on the ill condition of a system on the macro level. Chen Xiru proposed that [13]: if $0 < cond(N) < 100$, then N is not of ill condition; if $100 < cond(N) < 1000$, then N has moderate-degree ill condition; if $cond(N) > 1000$, then N has serious ill condition. But this measure index will be different to different problems. The eigenvalue-analysis method focuses on the several smaller eigenvalues. Suppose $\lambda_1, \lambda_2 \cdots \lambda_n$ $(\lambda_1 > \lambda_2 > \cdots > \lambda_n)$ are the eigenvalues of a matrix, if several eigenvalues are very small, it can be thought that there are some certain multicollinearities in the observation equation. From (28.8) we can see that for a symmetrical positive determined matrix, the condition-number method can also be considered to be an eigenvalue-analysis method. The VIF method is mainly concerned about the collinear relationship between the parameters and the parameters. The determinant-discriminant method is similar to the eigenvalue analysis. This method considers that: if determinant value of the matrix is close to zero, there will be approximate linear relationship between rows or columns. The value of the determinant is the product of the eigenvalues so it can be considered that the determinant-discrimination method is the extension of the eigenvalue-analysis method.

From the above analysis we can see that the several methods are mainly concerned about the macroscopic ill condition of a system or the multicollinearities between parameters. It's lack of concern for the inherent and detailed laws of the system. Compared the above methods, matrix spectral analysis method [5, 13] used in this paper can be more comprehensive and intuitive to reflect the inherent variation and laws of the system in the calculation process. The "spectrum" of a linear operator is defined as the collection of the eigenvalues of the operator matrix. Through this method, we can focus on not only the maximum or the minimum eigenvalue but also the change of every eigenvalue. It transfers the scattered and partial data to continuous line pattern, so we can see the results of the trend more fully and more intuitively.

28.4 Analysis of Ill Posedness in Multiple GNSS System Carrier High-Precision Double-Differential Positioning

The running angular velocity of GPS, GLONASS satellites and the IGSO and GEO satellites of Compass as shown below:

As can be seen from the above (Fig. 28.1), the running angular velocity of GLONASS satellites is the fastest . Compared to GLONASS, the GPS satellites run slightly more slowly. IGSO and GEO satellites of Compass have the slowest running angular velocity, especially for the GEO satellites which are the earth synchronous satellite and the running velocity angular of which is almost zero. This is consistent with their orbital altitude and operating characteristics.

With growth of the observation time, the geometric relationship between the satellite and the station changes much more in the process of ambiguity resolution. The changes of the eigenvalues of the normal matrix are shown as following figures (The results are shown by two modes: conventional baseline solution and network RTK solution, using 1 h of 1 Hz sampling data collected by 3S Centre of Southeast University on July 13, 2012.):

From Fig. 28.2 we can see the tendency of the eigenvalues' change in conventional baseline solution and network RTK solution in the three systems. With the growth of the observation time, the larger eigenvalues become more and more consistent, and at the same time, the smaller eigenvalues become larger and larger

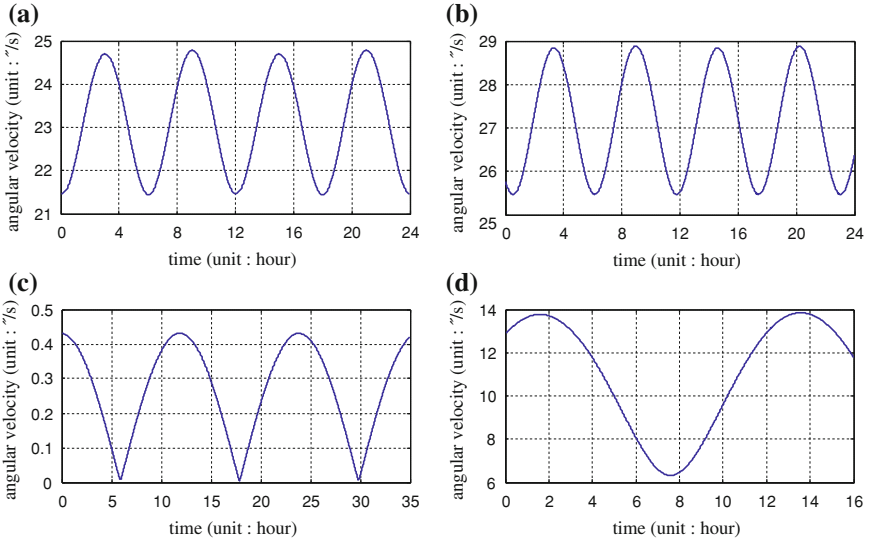


Fig. 28.1 The running angular velocity of satellites. **a** angular velocity of GPS satellites, **b** angular velocity of GLONASS satellites, **c** angular velocity of GEO satellites, **d** angular velocity of IGSO satellites

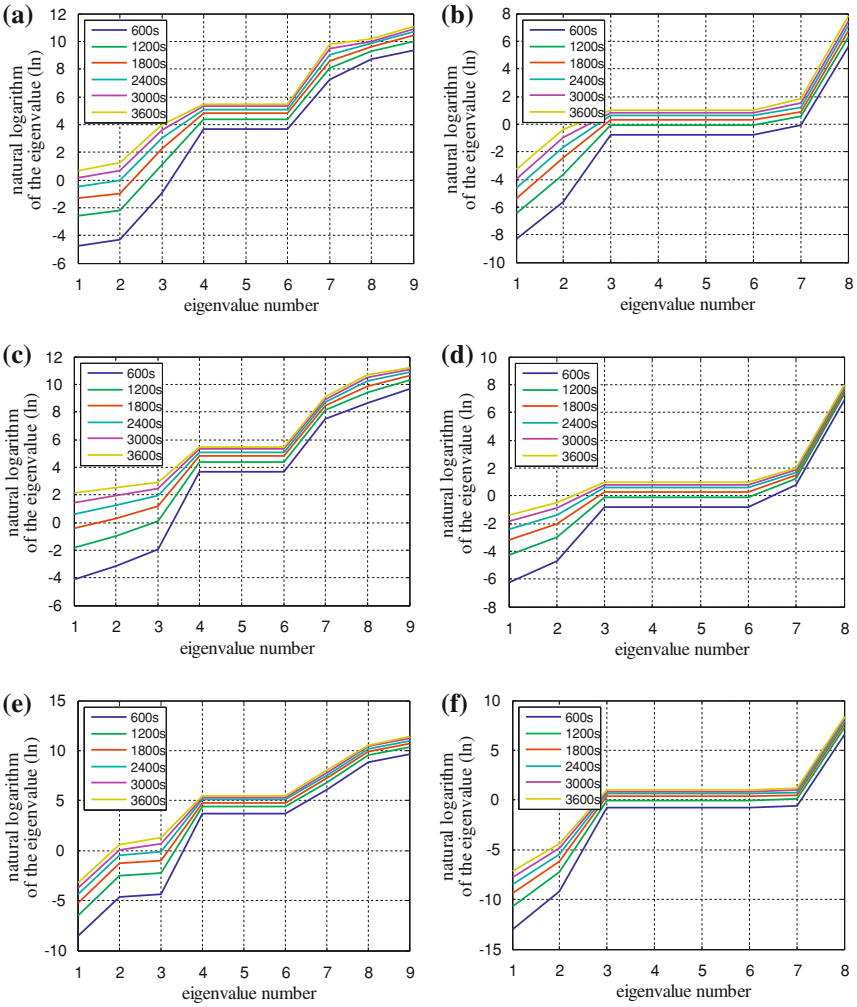


Fig. 28.2 The spectrums of eigenvalues of different time spans, **a** spectrum of eigenvalues in baseline solution of GPS, **b** spectrum of eigenvalues in network RTK of GPS, **c** spectrum of eigenvalues in baseline solution of GLONASS, **d** spectrum of eigenvalues in network RTK of GLONASS, **e** spectrum of eigenvalues in baseline solution of compass, **f** spectrum of eigenvalues in network RTK of compass

at faster speeds. So the ratio of the larger eigenvalues to the smaller eigenvalues becomes smaller. In other words, the eigenvalues are of convergence to closer and closer from each other. This also explains why the ill posedness of the normal matrix can be weakened with growth of the observation time from a mathematical point of view.

From the above analysis and Fig. 28.1, we can further speculate that if satellites have faster running angular velocity, the convergence of eigenvalues will be faster

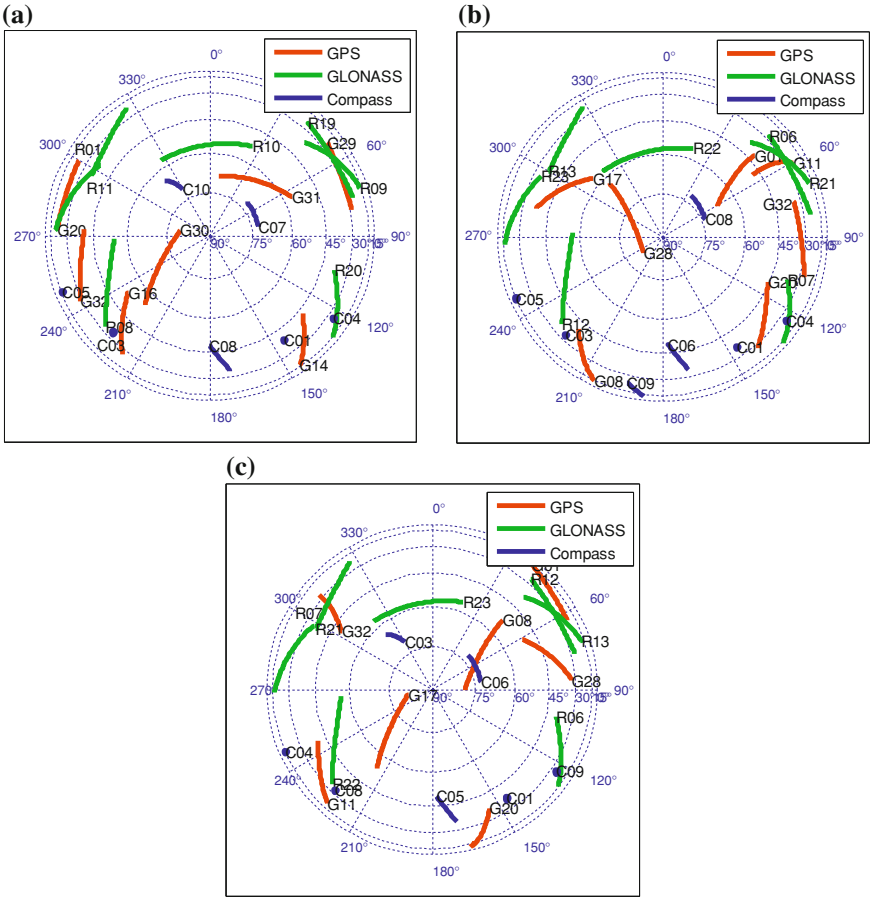


Fig. 28.3 The satellite distribution of three periods. **a** satellite distribution of 1th period, **b** satellite distribution of 2th period, **c** satellite distribution of 3th period

and the ill posedness will be weakened faster. We validate this inference by experiments. In the experiments, we use three-period observational data of three systems which are collected by 3S Center of Southeast University on July 13, 2012. Each period lasted for an hour and was at the sampling rate of 1 Hz. In each period, we all choose seven effective satellites for each system to try to avoid the influence of the number of satellites. The results are expressed with the form of contrasts. The satellite distribution of the three periods can be seen in Fig. 28.3 and the experimental results can be seen from Fig. 28.4 and Table 28.1:

It can be seen from Fig. 28.4a, c, e and Table 28.1, in conventional baseline solution, the larger eigenvalues of the normal matrix in GLONASS is closer to that in GPS. However, to the smaller eigenvalues especially the minimum eigenvalue, which in GLONASS is greater than that in GPS. This indicates that in conventional baseline solution process, the singularity of the normal matrix in GLONASS

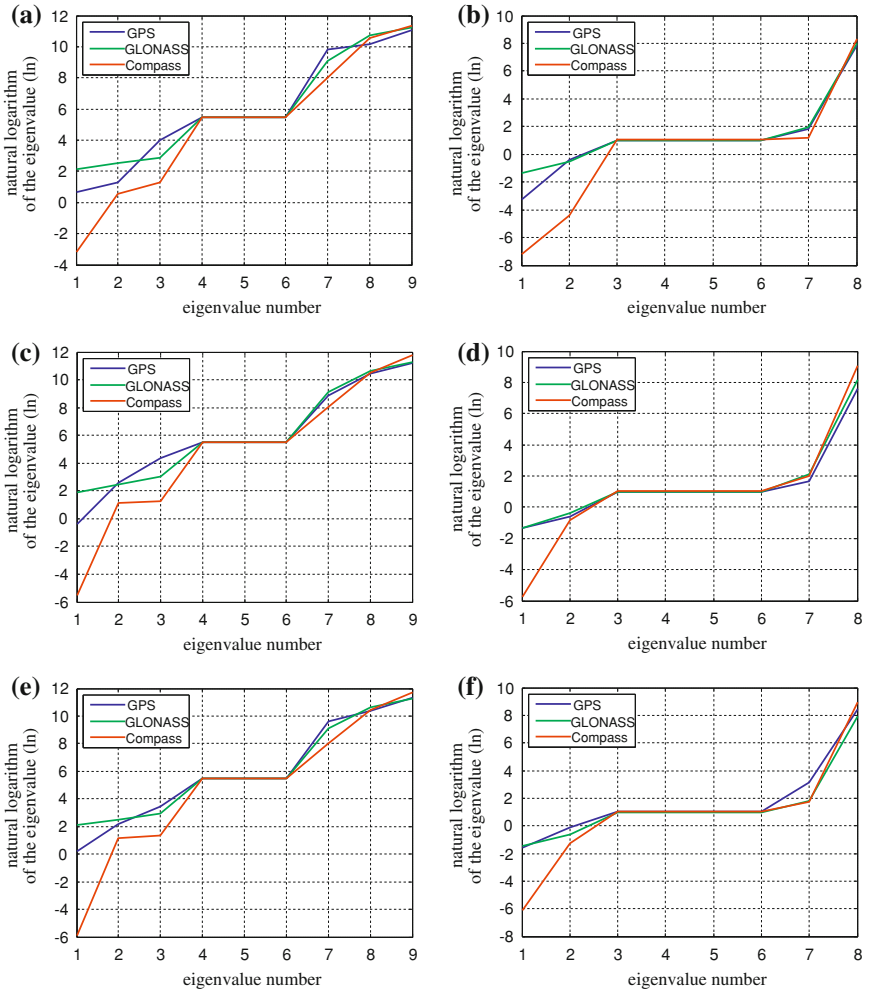


Fig. 28.4 The spectrums of eigenvalues of three periods, **a** spectrum of eigenvalues in baseline solution of 1th period, **b** spectrum of eigenvalues in network RTK of 1th period, **c** spectrum of eigenvalues in baseline solution of 2th period, **d** spectrum of eigenvalues in network RTK of 2th period, **e** spectrum of eigenvalues in baseline solution of 3th period, **f** spectrum of eigenvalues in network RTK of 3th period

is less than that GPS. In other words, the algebraic structure of GLONASS is slightly better than that of GPS in the process of conventional baseline solution. The larger eigenvalues of the normal matrix system in Compass is a bit larger than that in GPS and GLONASS, but its smaller eigenvalues especially the minimum eigenvalue are much smaller than that in GPS and GLONASS. Compared to GPS and GLONASS, Compass has the worst algebraic structure in conventional baseline solution. This is corresponding to (28.2). If satellites have the faster

Table 28.1 Condition number of the normal matrix in three periods

	Natural logarithm of condition number of the normal matrix (ln)					
	Conventional baseline solution			Network RTK solution		
	G	R	C	G	R	C
Period 1	10.404	9.092	14.576	11.118	9.388	15.513
Period 2	11.605	9.396	17.312	8.924	9.511	14.768
Period 3	11.120	9.130	17.623	10.026	9.362	15.059

Note "G" "R" "C" in the table represents GPS, GLONASS and compass separately)

running angular velocities, the geometric structure between satellites and stations changes more between epochs and the coefficients of observations equations also change faster, so the correlation between the equations is weakened faster.

In Network RTK solution, the elevation angles of a satellite to two stations are very close generally. The coefficients of wet-troposphere delay of two stations are quite close to each other. This results in that not only equations between adjacent epochs have strong correlation but also there are strong multicollinearities between parameters, which make the ill posedness more seriously. The height of satellite orbit and the running angular velocity of GPS are close to that of GLONASS. In addition, due to other factors such as satellite distribution and the choice of reference satellite, the difference of the algebraic structure between GPS and GLONASS in Network RTK is not as obvious as that in conventional baseline solution. But from Fig. 28.4b, d, f and Table 28.1 can be seen, GPS or GLONASS is still of much difference with Compass. Particularly in the minimum eigenvalue, GPS or GLONASS is much larger than Compass. From the above analysis, we can get the conclusion that the running angular velocities of satellites play a major role among the factors related to ill-posedness problem.

28.5 Conclusions

In this paper, a method called analysis of characteristic spectral of a matrix is represented to analyze the algebraic structure in GNSS high-precision differential positioning comparatively. The method describes the characteristics more fully and more intuitively. In the process of GNSS high-precision differential positioning, the properties of the ill posedness are consistent with the running characteristics of satellites. The faster satellites run, the more quickly the geometric structure that consists of satellites and stations changes, and the faster the ill posedness is weakened. This is obvious in conventional baseline solution particularly. Due to the existence of high orbit satellites especially the GEOs, the ill posedness in the process of GNSS high-precision differential positioning of Compass is much serious than that of GPS or GLONASS at the present stage.

Acknowledgments This work is supported by the Open Research Fund of The Academy of Satellite Application under grant No. 2012-1514.

References

1. Yang YX (2010) Progress, contribution and challenges of compass/beidou satellite navigation system. Acta Geodaetica et Cartographica Sinica 39(1):1–6
2. Gao CG, Hu W (2011) The principle and application of satellite navigation and positioning. China Communications Press, Beijing, pp 14–16, 153–156
3. Yang XC, Xu B, Hu Y (2012) Performance research of constellation for COMPASS . http://www.cnki.net/kcms/detail/11.4415.P.20121023.0846.004.html
4. Guo Q, Hu Z (2007) The effect of ill-condition equations of GPS rapid positioning on the GPS baseline solution. Sci Surveying Mapp 2:103
5. Teunissen P, de Jonge P, Tibrius C (1994) On the spectrum of the GPS DD-Ambiguities. Proceedings ION GPS-94, Salt Lake City, Utah, USA, September 20–23, 1994
6. Xu P, Shen Y, Fukuda Y, Liu Y (2006) Variance component estimation in linear inverse ill-posed models. J Geod 80:69–81
7. Hao W, Li F (2010) A new method for ill-conditioned diagnosis based on spatial analysis. J Inf Comput Sci 7(9):1846–1853
8. Wang Zhengjie, OU Jikun (2004) A new kind of correction scheme to the singular value in the ill-posed problems and its application in the measurement. 14(6):672–676
9. WANG SL, Wang Q, Pan SG, Yu XX (2011) New cycle-slip detection and repair method for CORS network. J Chinese Inertial Technol 9(5):590–593
10. Huang DF, Xiong YL, Yuan LG (2006) Global positioning system (GPS)-theory and practice. Xi'an Jiaotong University Press, Chengdu, pp 71–73
11. Sheng XF (2011) Algorithm for network RTK based on star structure and precise satellite clock error estimation. Transportation College of Southeast University, Nanjing
12. Niell AE (1996) Global mapping functions for the atmosphere delay at radio wavelengths. J Geophys Res 101:3227–3246
13. Lu XS, Feng Z, Liu J (2007) Analysis theory on ill-conditioned system with application in surveying. Surveying and Mapping Press, Beijing, pp 7–15

Chapter 29
GPS/GLONASS System Bias Estimation and Application in GPS/GLONASS Combined Positioning

Junping Chen, Pei Xiao, Yize Zhang and Bin Wu

Abstract Multi-GNSS data analysis has become a new challenge with the development of satellite navigation systems. System bias is the key issue in Multi-GNSS data analysis, which has no recommended models within IGS community. We introduce the integrated data analysis model developed at the GNSS data analysis center of Shanghai Astronomical Observatory (SHAO). Based on the routine GNSS data analysis at SHAO over 14 months, we analyze the precise GPS/GLONASS system bias product in detail. Results show: (1) system bias shows similarity for same type of receivers, while obvious difference are observed for different type of receivers; (2) variation of system bias shows same pattern for all stations, which indicates that the long-term variation of system bias is caused by the system time offset; (3) system bias is influenced also by type of antenna type. A model is derived to separate hardware delay difference (HDD) between GPS/GLONASS observations at the same receiver and the so-called inter-frequency bias (IFB). Analysis of the HDD and IFB time series shows that both terms are affected by the change of receiver type, antenna type, firmware series, cable type and length. Applying the system bias into PPP positioning, precision of GLONASS-only solution is improved by 55 % and precision of GPS/GLONASS combined solution is improved by 30 %.

Keywords GNSS · SHA · Analysis center · Inter system bias (ISB) · IFB

J. Chen (✉) · P. Xiao · Y. Zhang · B. Wu
Shanghai Astronomical Observatory, Chinese Academy of Science, Shanghai, People's Republic of China
e-mail: junping@shao.ac.cn

P. Xiao · Y. Zhang
College of Surveying and Geo-Informatics, Tongji University, Shanghai, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_29,

29.1 Introduction

Coordinate and time reference frame are both the key parameters of satellite navigation system. As to time reference frame, GPS is based on GPST, GLONASS is based on GLONASST. As to coordinate reference, GPS adopts WGS-84, GLONASS adopts PZ-90. There are differences in framework accuracy and scale for different navigation systems [1–4].

As navigation system develops and updates, multi-system fusion has become the tendency of the development. In certain environment, such as urban canyon and ravines, single system can't provide service because of limited satellite conditions. Besides, satellite constellation has periodic regression relative to the Earth, the relative relationship of navigation satellites–Earth–Sun also has regression of different period. The periodic regression of these relative relationships will add relevant periodic errors into parameters such as coordinates and receiver clock offset [5]. So multi-mode observation increases the number of available satellites. Meanwhile, the data fusion can reduce the influence of the periodic regression of satellite constellation, to improve the precision of station-relative parameters (e.g. coordinate, troposphere) and other public parameters (e.g. ERP) [6].

High-precision integrated processing of multi-mode data is the guarantee of multi-system fusion. Multi-system integrated data processing need taking into consideration of various system bias parameters. As to GPS/GLONASS integrated processing, the inter-system bias (ISB) of GPS and GLONASS system includes system's time difference TO and different system's hardware delay bias difference in receiver (ΔDCB). Among them, system's time difference is the difference between system times; ΔDCB is the hardware delay difference in receiver, it includes inter-frequency bias (IFB) of GLONASS satellite, which is because GLONASS system is based on frequency division multiple access (FDMA). Nowadays, IGS has not published integrated product and its processing standard. This article introduces GPS/GLONASS integrated data processing model; analyses and discusses the characteristics of ISB and IFB parameter, which is based on the 14 months routine results of global GNSS network provided by the GNSS data analysis center at SHAO (SHA). By introducing bias parameters to multi-mode positioning, the parameter resolution precision can be greatly improved.

29.2 GPS/GLONASS Integrated Data Processing Unified Model

Observation function of receiver i to GPS satellite j can be written as:

$$\begin{aligned} P_{\mathrm{i}}^{\mathrm{j}} &= \rho_{\mathrm{i}}^{\mathrm{j}} + \mathrm{c} \cdot \left(dt_i - dt^j\right) + DCB_i^j - I_i^j + T_i^j + \varsigma_i^j \\ \mathrm{L}_{\mathrm{i}}^{\mathrm{j}} &= \rho_{\mathrm{i}}^{\mathrm{j}} + \mathrm{c} \cdot \left(dt_i - dt^j\right) + DPB_i^j + \lambda \cdot N_i^j - I_i^j + T_i^j + \varepsilon_i^j \end{aligned} \tag{29.1}$$

P_i^j, L_i^j are respectively the pseudorange and carrier phase observation; ρ_i^j is geometrical distance; c is light speed, λ is wavelength; dt_i is receiver clock offset, dt^j is satellite clock offset; DCB_i^j and DPB_i^j are pseudorange and carrier phase bias (including both receiver and satellite); N_i^j is ambiguity, I_i^j is the ionospheric delay error, T_i^j is the tropospheric delay error, ς_i^j is other error corrections (including relativistic effect, tide, PCO, PCV, phase unwrapping and so on), ε_i^j is residual. In practical application, I_i^j could be ignored by forming the ionosphere-free combination using dual-frequency pseudo-range and carrier phase observations.

The pseudorange observation in (29.1) provides reference to clock offset parameters. The pseudorange bias DCB_i^j (such as P1-P2, P1-C1) will be absorbed by clock offset $\mathrm{c} \cdot (dt_i - dt^j)$. Nowadays, the carrier phase bias DCB_i^j is not included in GPS data processing, it will be combined with other parameters (mainly ambiguity). Then (29.1) can be rewritten as:

$$\begin{aligned} P_{\mathrm{i}}^{\mathrm{j}} &= \rho_{\mathrm{i}}^{\mathrm{j}} + \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^j\right) - I_i^j + T_i^j + \varsigma_i^j \\ \mathrm{L}_{\mathrm{i}}^{\mathrm{j}} &= \rho_{\mathrm{i}}^{\mathrm{j}} + \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^j\right) + \lambda \cdot \overline{N}_i^j - I_i^j + T_i^j + \varepsilon_i^j \end{aligned} \tag{29.2}$$

where:

$$\begin{aligned} c \cdot \left(\overline{dt}_i - \overline{dt}^j\right) &= c \cdot \left(dt_i - dt^j\right) + DCB_i^j \\ \lambda \cdot \overline{N}_i^j &= \lambda \cdot N_i^j + DPB_i^j - DCB_i^j \end{aligned} \tag{29.3}$$

Nowadays, IGS clock offset product reference is based on P1/P2 ionosphere-free combination. Under that basis, the DCB_i^j of P1/P2 ionosphere-free combination in (29.3) will be absorbed by clock offset parameter. Other observations need parameters provided by IGS to correct DCB_i^j.

Expend (29.2) to GPS/GLONASS dual-mode observation, observation function of receiver i to GPS satellite k and GLONASS satellite j can be written as :

$$\begin{aligned} \mathrm{L}_{\mathrm{i}}^{kG} &= \rho_{\mathrm{i}}^{kG} + \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^k\right)^G - I_i^{kG} + T_i^{kG} + \lambda^G \cdot \overline{N}_i^{kG} + \varsigma_i^k \\ \mathrm{L}_{\mathrm{i}}^{jR} &= \rho_{\mathrm{i}}^{jR} + \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^j\right)^G + ISB_i^{jk} + \lambda^R \cdot \overline{N}_i^{jR} - I_i^{jR} + T_i^{jR} + \varepsilon_i^j \end{aligned} \tag{29.4}$$

where:

$$\begin{aligned} ISB_i^{jk} &= \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^j\right)^R - \mathrm{c} \cdot \left(\overline{dt}_i - \overline{dt}^k\right)^G \\ &= TO + \Delta DCB_i^{j,k} \end{aligned} \tag{29.5}$$

In (29.4), the superscript R represents GLONASS, G represents GPS; ISB_i^{jk} is inter-system bias on station i between GPS satellite k and GLONASS satellite j (including system time difference TO and pseudorange delay bias in satellites and receiver $\Delta DCB_i^{j,k}$, which includes inter-frequency bias IFB_i^j). Definitions of other parameters are the same with (29.1) and (29.2). TO in (29.5) is defined as a one-

day constant for all stations. $\Delta DCB_i^{j,k}$ is mainly because of GPS and GLONASS systems' different frequencies, which can be written as ΔDCB_{sys}. $\Delta DCB_i^{j,k}$ is also slightly effected by GLONASS satellites' different frequencies IFB_i^j, IFB_i^j is various to different receivers and different frequencies.

Formula (29.4) is the universal observation function of multi-system integrated data processing, it also applies to the combined observation of GPS and other satellite system. By defining inter-system bias ISB_i^{jk}, estimating $c \cdot \left(\overline{d}t_i - \overline{d}t^j\right)^G$ and unifing GLONASS clock offset to GPS time system, we can realize the integrate of multi-system's time reference. ρ_i^j Contains satellite orbit and receiver coordinates, restrain station coordinates to ITRF reference, then we can realize the unification of multi-system's space reference. As to users, adopting these orbit and clock offset and all kinds of bias parameters under the same time and space reference can unifies different systems' observation to the same satellite system, thus simplify the users' application and promote positioning precision.

In (29.4), calculating $\overline{d}t_i$ and $\overline{d}t^j$ at the same time is rank deficient, the general method is fixing one reference clock (usually the station with an external high-precision atomic clock, fixing the clock offset by GPS pseudorange process). ISB_i^{jk} contains TO, ΔDCB_{sys}, and IFB_i^j, IFB_i^jcan be absorbed in $\lambda^R \cdot \overline{N}_i^{jR}$ while TO and ΔDCB_{sys} have correlations with clock offset. Considering these correlations above, there are two solutions: weight ISB_i^{jk} to reduce the influence on correlation; add zero mean condition to all the ISB_i^{jk} in one station (IGS AC Mail 643). Different solutions cause the inconformity of GLONASS clock offset reference between IGS analysis centers [7].

29.3 GPS/GLONASS Inter-System Bias

Based on the multi-system integrated data processing modal above, Shanghai Astronomical Observatory developed integrated Geodetic Platform Of SHAO (iGPOS) and established GNSS data analysis center (SHA) [8].

Figure 29.1 shows the IGS network processed in the GNSS routine of SHA, among them about 70 stations can provide GPS/GLONASS combined observations. Figure 29.2 compares several analysis centers' orbit precision (from 2011.7 to 2012.8). The precision of GPS orbits of SHA is 1.5 cm and the precision of GLONASS orbits is 3.2 cm, which is about the precision of other IGS analysis centers.

SHA adopts the strategy that EMR and GFZ uses to deal with ISB: set the ISB of each receiver to each GLONASS frequency as a one-day constant. Figure 29.3 shows the GPS/GLONASS ISBs of station BRMU (BERMUDA, UK) from 2011181 to 2012240. In this figure, different color represents different GLONASS satellite frequency. In this period ISBs are between 50 and 70 m, the difference of

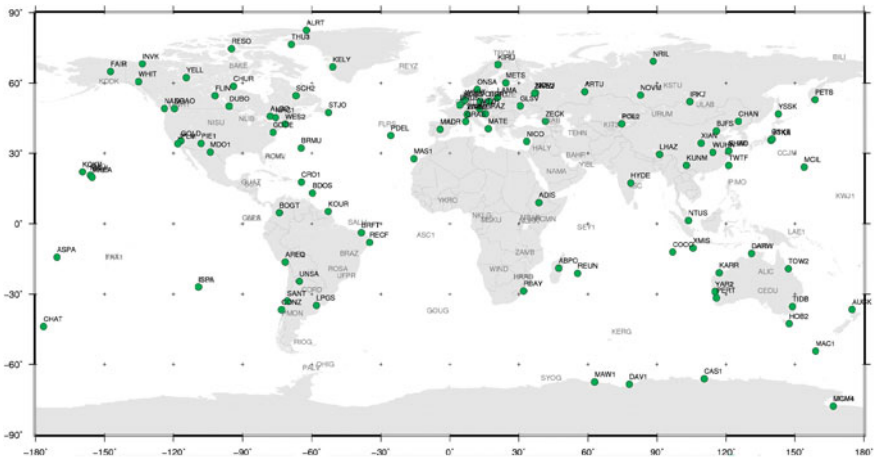


Fig. 29.1 IGS network processed in the GNSS routine of SHA

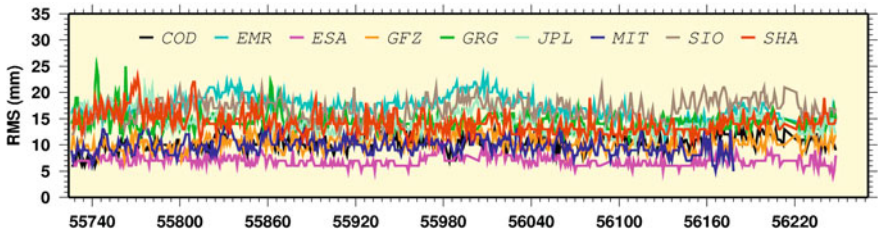


Fig. 29.2 Comparison of IGS analysis centers' orbit precision. Results in mm

adjacent day is less than 3 ns. Difference of different GLONASS frequency is less than 5 m (the minus channel number is −7, the max is 6). IFB's order of magnitude is obviously lower than ISB. Besides, on 2011271 BRMU's antenna type changed from TRM29659.0 to JAVRINGANT_DM, this change reflected to ISB obviously (about 10 m). It can be concluded that the type of antenna has influence on ISB.

Figure 29.4 shows the ISB series of 26 LEICA receivers. Different color represents different antenna type. It can be seen that the ISB of stations with LEICA antenna (LEIAT504GG and LEIAR25.R3), Topcon antenna (TPSCR3_GGD), Allen Osborne antenna (AOAD/M_T and AOAD/M_B) and Javad antenna (JAVRINGANT_DM) only have little difference less than 5 m. Meanwhile Ashtech and AOAD/M_TA_NGS antenna (this kind of antenna adopts Ashtech low noise amplifier technology [9]) and Trimble antenna (TRM29659.00) have obviously bigger difference. However, the difference between different antenna types is relatively less than the difference between receiver types.

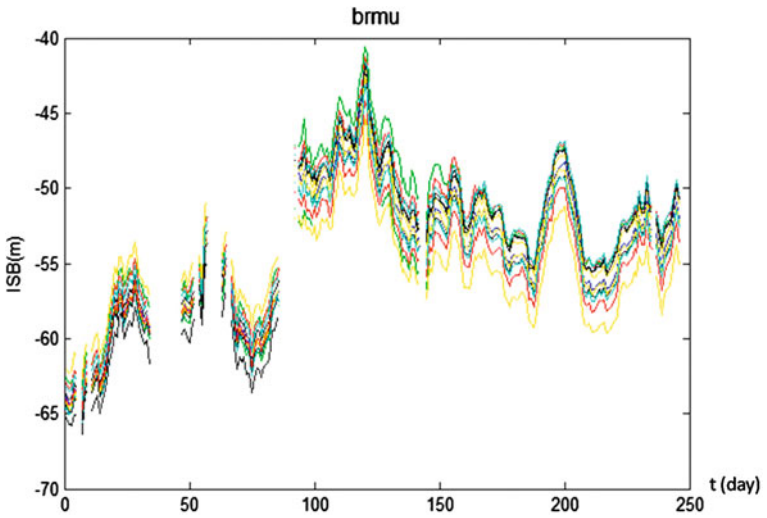


Fig. 29.3 ISBs of different GLONASS satellites on station BRMU (2011.06.30–2012.8.30)

As mentioned, *ISB* includes 3 parts: system time difference *TO*, GPS and GLONASS systems' hardware delay bias difference ΔDCB_{sys}, and GLONASS satellites' inter-frequency bias IFB_i^j. As it shows in Fig. 29.4, ISB's long-term changing tendency is consistent for the same type receivers, it mainly reflects the long-term changes of *TO* and ΔDCB_{sys}. Taking one frequency as reference frequency (such as channel 0) can reduce hardware delay bias difference in station and system time difference:

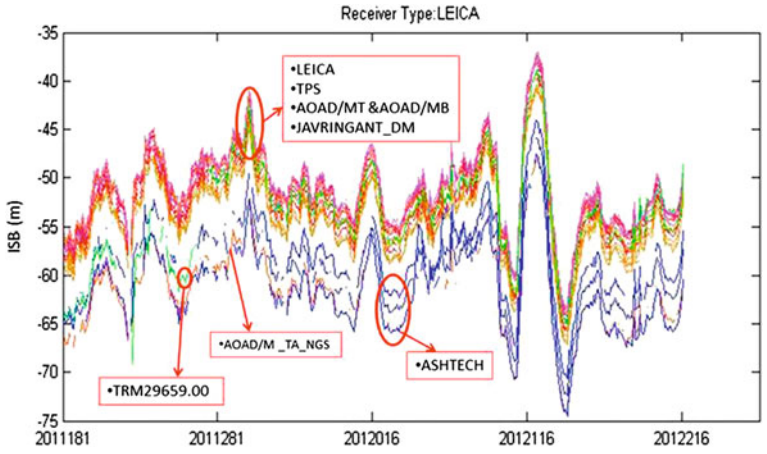


Fig. 29.4 ISB series of 26 LEICA receivers (2011.06.30–2012.8.30)

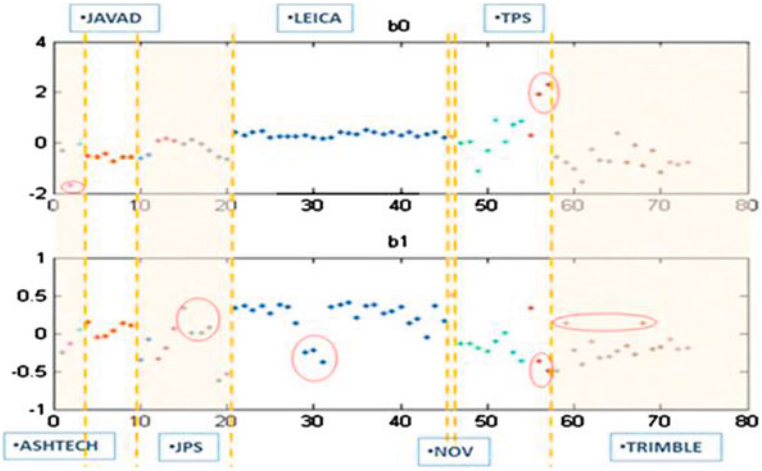


Fig. 29.5 $b0$, $b1$ of all the 74 stations

$$\begin{aligned} ISB_i^m - ISB_i^n &= \left(TO + \Delta DCB_i^{m,G}\right) - \left(TO + \Delta DCB_i^{m,G}\right) \\ &= \left(TO + \Delta DCB_{sys} + IFB_i^m\right) - \left(TO + \Delta DCB_{sys} + IFB_i^n\right) \\ &= IFB_i^{m,n} \end{aligned} \tag{29.6}$$

IFB has linear relationship with the channel number [10, 11], so (29.6) can be rewritten as:

$$ISB_i^m - ISB_i^n = IFB_i^{m,n} = b0 + b1 \cdot (f^m - f^n) \tag{29.7}$$

In (29.7), f^m, f^n are GLONASS satellites' channel numble, $b0$, $b1$ are the fitting coefficients.

Take channel 0 as reference frequency, subtract its ISB from other satellite's. By means of the least square fit according to formula (29.7), we obtain $b0$, $b1$ for each station on each day. By using 14 months ISB data, which is provided by Shanghai Observatory GNSS Analysis Center (SHA), we get all the $b0$, $b1$ of 74 stations (shown in Fig. 29.5). There are 7 receiver types in total, it is shown that $b0$, $b1$ values of the same type receivers are consistent, and $b0$, $b1$ of different receivers vary widely. Antenna type's influence on $b0$, $b1$ is also obvious, the 11 stations with Ashtech antenna has been marked by red circles in Fig. 29.5, $b0$, $b1$ of these stations have obvious difference.

29.4 Application of ISB in GPS/GLONASS Combined Positioning

Introducing ISB to positioning can improve the accuracy and validity of positioning, especially when the valid satellite number is less [12]. ISB can be corrected directly, then GPS and GLONASS can be seen as a single system.

GPS/GLONASS combined positioning function is:

$$\begin{aligned} \mathrm{P}_{\mathrm{i}}^{kG} &= \rho_{\mathrm{i}}^{kG} + \mathrm{c}\cdot\overline{d}t_i - I_i^{kG} + T_i^{kG} + \varsigma_i^k \\ \mathrm{P}_{\mathrm{i}}^{\mathrm{j}R} &= \rho_{\mathrm{i}}^{\mathrm{j}R} + \mathrm{c}\cdot\overline{d}t_i + ISB - I_i^{jR} + T_i^{jR} + \varepsilon_i^j \\ \mathrm{L}_{\mathrm{i}}^{\mathrm{j}G} &= \rho_{\mathrm{i}}^{kG} + \mathrm{c}\cdot\overline{d}t_i + \lambda^G\cdot\overline{N}_i^{kG} - I_i^{kG} + T_i^{kG} + \varsigma_i^k \\ \mathrm{L}_{\mathrm{i}}^{\mathrm{j}R} &= \rho_{\mathrm{i}}^{\mathrm{j}R} + \mathrm{c}\cdot\overline{d}t_i + ISB + \lambda^R\cdot\overline{N}_i^{jR} - I_i^{jR} + T_i^{jR} + \varepsilon_i^j \end{aligned} \tag{29.8}$$

The parameters in formula (29.8) are the same with formula (29.4), GPS/GLONASS satellite clock offset adopts SHA's precision product, ISB can be obtained by using the model above mentioned. To the station with known ISB, there are only 6 parameters (coordinates, receiver clock offset and troposphere parameter) to be estimated. To the station with unknown ISB, we can give ISB an initial value according to the receiver type and antenna type, use the IFB model above to correct ISB, we only need another parameter (ISB of channel 0), then we can reduce the number of parameters and promote positioning precision.

29.4.1 Pseudorange Positioning

We choose 4 stations' data (pots, casl, chur, aspa) on doy 120 to doy 126, 2012, the interval is 30 s. These stations are installed with different manufacturers' receivers, the receiver and antenna information of these stations is in Table 29.1.

Tests are conducted by using pseudorange observations in 2 strategies: GONASS only and GPS/GLONASS combined positioning. Every strategy is applied in 3 scenarios:

1. Without consideration of GLONASS IFB;
2. Introduce GLONASS IFB from SHA;
3. Introduce GLONASS IFB modle of the corresponding receiver; estimate a one-day parameter: ISB of frequency-0.

The coordinates precision and increase rate are shown in Tables 29.2 and 29.3.

It can be seen from the statistics in Tables 29.2 and 29.3 that without considering GLONASS IFB obtains the lowest precision, while directly using ISB provided by SHA obtains the highest precision. The two ways to introduce IFB both greatly improve the positioning precision, coordinate precision increases up to

Table 29.1 Station information

Station	Receiver type	Antenna type
pots	JAVAD TRE_G3TH DELTA	JAV_RINGANT_G3T
cas1	LEICA GRX1200GGPRO	AOAD/M_T
chur	TPS NET-G3A	ASH701945E_M
aspa	TRIMBLE NETR5	TRM55971.00

Table 29.2 GLONASS pseudorange positioning coordinates precision and increase rate

Station	Without IFB RMS (m)	IFB model		ISB from SHA	
		RMS (m)	Increase rate (%)	RMS (m)	Increase rate (%)
chur	7.60	5.84	23.19	3.40	55.28
aspa	4.29	3.60	16.09	3.30	22.99
cas1	2.99	1.90	36.29	1.77	40.87
pots	2.39	2.29	4.11	1.89	20.79

Table 29.3 GPS/GLONASS combined pseudorange positioning coordinates precision and increase rate

Station	Without IFB RMS (m)	IFB model		ISB from SHA	
		RMS (m)	Increase rate (%)	RMS (m)	提高率 (%)
chur	2.99	2.42	19.17	1.89	36.75
aspa	2.62	2.41	8.18	2.31	11.85
cas1	2.34	1.75	25.12	1.65	29.56
pots	1.74	1.71	1.43	1.58	9.08

55 %. Precision of combined positioning is up to 4 times better than of GLONASS single system (chur, from 7.60 to 1.89 m).

29.4.2 Carrier Phase Positioning

We make a test of Kinematic PPP with the carrier phase observation data at the station CHUR on 2012 doy 318. This test is applied in 4 scenarios:

1. GPS PPP
2. GLONASS PPP
3. Combined GPS/CLONASS PPP
4. Based on the third strategy, introduce the inter-system hardware delay bias IFB, which is provided by Shanghai Observatory GNSS Analysis Center.

All of these four strategies obtain satisfactory final positioning results. Figure 29.6 shows the positioning error and its components in X, Y, Z directions on the first 50 epochs. It can be seen that the convergence speeds of these four strategies are all

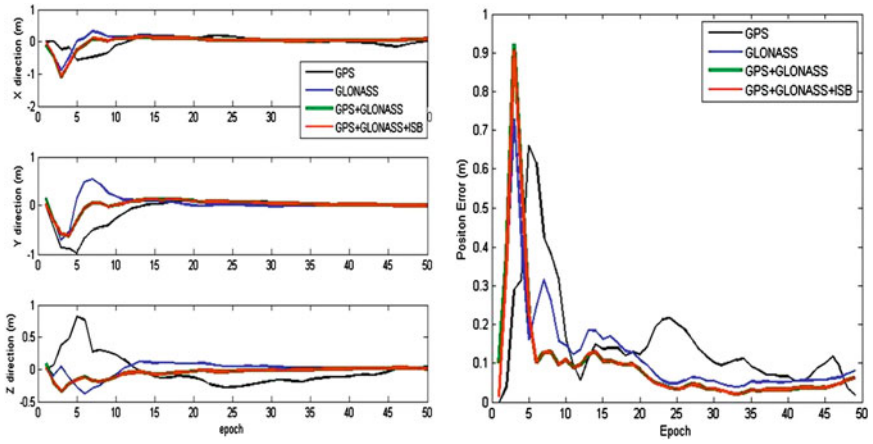


Fig. 29.6 Kinematic carrier phase PPP results on station CHUR

fast, among them the convergence of combined GPS/GLONASS PPP is faster than single system PPP. Introducing ISB only has little effect on the first several epochs. The results of strategy (29.3) and (29.4) are almost the same, this is because in strategy (29.3), the ISB values have been absorbed by ambiguities, although it gets good precision and fast convergence speed, its ambiguities are not accurate.

29.5 Conclusion

As navigation system develops and updates, multi-system fusion has become the tendency of the development. Multi-mode can solve the issues of coverage and system bias that may exist in single system. In certain environment, such as urban canyon and ravines, single system could not provide service because of limited satellite conditions. We introduce the integrated data analysis model by GNSS data analysis center of Shanghai Astronomical Observatory (SHAO). Based on the routine GNSS data analysis at SHAO over 14 months, we analyze the precise GPS/GLONASS system bias product in detail and put forward a model of inter-frequency bias (IFB). Results show: (1) ISBs have similarity in receivers of the same type, while obvious difference are observed in receivers of different type; (2) variation of ISBs shows same pattern for all stations, which indicates that the long-term variation of ISB is caused by the system time offset; (3) ISB is influenced also by antenna type. Applying ISB into pseudorange positioning, precision of GLONASS-only solution is improved by 55 % and precision of GPS/GLONASS combined solution is improved by 30 %. Applying ISB into PPP, there is no obvious effect on coordinates, this is because in PPP, ISB can be absorbed by ambiguity.

Acknowledgments This paper is supported by the 100 Talents Programme of The Chinese Academy of Sciences, the National High Technology Research and Development Program of China (Grant No. 2013AA122402), and the National Natural Science Foundation of China (NSFC) (Grant No. 40974018 and 11273046).

References

1. Brown K (1991) The theory of the GPS composed clock. In: Proceedings of ION GPS-91, NM, Institute of Navigation, Albuquerque, pp 223–241
2. GLONASS Interface Control Document (2008): Edn 5.1, Russian Institute of Space Device Engineering
3. Delporte J (2009) The definition and implementation of Galileo system time (GST), ICG-4 WG-D on GNSS time scales
4. Wei Z (2010) The time and space reference in satellite navigation system. China satellite navigation conference (CSNC), Beijing
5. Flohrer C (2008) Mutual validation of satellite-geodetic techniques and its impact on GNSS orbit modeling. Geodaetisch-geophysikalischeArbeiten in der Schweiz, vol 75
6. Dach R, Schaer S, Lutz S, Meindl M, Beutler G (2010) Combining the observations from different GNSS, EUREF 2010 Symposium, June 02–05, 2010, Gävle, Sweden
7. Dach R, Schaer S, Meindl M (2012) Comparison of GPS/GLONASS clock solutions, IGS–workshop on GNSS–Biases, Bern, Switzerland, Jan 18–19, 2012
8. Chen J, Wu B, Hu X, Li H (2012) SHA: the GNSS analysis center at SHAO. In: Lecture notes in electrical engineering (LNEE), vol 160, pp 213–221
9. http://www.ngs.noaa.gov/ANTCAL/Antennas.jsp?manu=Allen+Osborne
10. Wanninger L (2012) Carrier-phase inter-frequency biases of GLONASS receivers. J Geod 86 (2):139–148. doi:10.1007/s00190-011-0502-y
11. Al-Shaery A, Zhang S, Rizos C (2012), An enhanced calibration method of GLONASS inter-channel bias for GNSS RTK. GPS Solut (online-first), DOI 10.1007/s10291-012-0269-5
12. Pei X, Chen J, Wang J, Zhang Y, Li H (2012) Application of Inter-system Hardware Delay Bias in GPS/GLONASS PPP. Lect Notes Electr Eng 160(2):381–387. doi:10.1007/978-3-642-29175-3_34

Chapter 30
Analysis of Multipath Effect on the Antennas of CORS Receivers

Shengjia Tang, Hui Liu and Shujing He

Abstract Though state-of-the-art Trimble dual-frequency receivers are employed in Continuously Operating Reference Stations (CORS) in China [1], the accuracy of CORS coordinate are affected by multipath, receiver noise, etc., while multipath is generally considered as the main contributor. To make a wisely choice on receivers and antennas in the construction of CORS, Trimble NetR5 with TRM55971.00 antenna and Trimble R8 Model 2 with TRM60158.00 antenna, which are common used, are selected to making a comparison in terms of multipath effect. The same data is calculated by TEQC for the sake of validating the algorithm this paper used. The answer shows that the multipath effect of Trimble NetR5 is better than R8 Model 2 about 12 % in the same place in adjacent days. What's more, to ensure that Trimble NetR5 has a stable performance, the multipath effect time series of an IGS station is calculated.

Keywords Multipath effect · CORS · TEQC · Antenna · Trimble

This paper was funded by the National high technology research and development plan (863 Program), 2007AA12Z309; The National Basic Research Program of China (973 Program) (2007CB310805).

S. Tang (✉) · H. Liu
GNSS Research Center, Wuhan University, No.129, Luoyu Road, Wuhan 430079, China
e-mail: tangsj114@gmail.com

H. Liu
e-mail: loweliu@hotmail.com

S. He
Fujian Institute of Surveying and Mapping, Fuzhou, Fujian Province 350003, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_30,

30.1 Introduction

Multipath effects are inevitable for the reflections from ground and surfaces of other objects [2]. Moreover, due to complicated conditions of observation, there are no universal effective algorithms in eliminating multipath effects. To decrease the effects of multipath, and increase the quality of observation data, 10 or more degree of elevation mask and anti-multipath antennas technique are usually adopted [3].

Trimble NetR5 GNSS infrastructure receiver with TRM55971.00 antenna, designed as a permanent reference station for use in Trimble VRS network, is common employed in CORS in China, such as GDCORS, JXCORS. And Trimble R8 Model 2 with TRM60158.00 antenna is an integrated system and widely used in GNSS project. Both receivers can track all GPS signals (L1/L2/L5) as well as GLONASS(L1/L2). By finishing the pre-treatment of observation files of 2 receivers, and contrasting the result TEQC calculated, an overall analysis is made including the relationship between multipath value and satellite elevation, multipath effects on different types of antennas.

30.2 Preliminaries

30.2.1 Cycle Slip Detection

Before calculate the multipath value of each epoch, observation data must be cleaned by detecting cycle slips and deleting the gross error and bad epoch. A method integrating M-W combination and ionosphere residual combination, called TurboEdit [4], is used in this paper.

Without consideration of atmospheric biases, multipath, receiver noise and high order of ionosphere error, the distance formula of phase and pseudorange are as follows:

$$\begin{aligned} \lambda_{\mathrm{i}}\varphi_{\mathrm{i}} &= \rho - \frac{I}{f_i^2} - \lambda_i N_i \\ P_i &= \rho + \frac{I}{f_i^2} \quad (i = 1, 2) \end{aligned} \tag{30.1}$$

where ρ is the distance between the satellite and the site, and I/f_i^2 is ionospheric delay. Combining the equations, it is easily to get the M-W combination which has a longer wave length about 86 cm, and can detect cycle slips in most conditions, and ionosphere residual combination which eliminate the orbit error, clock error of receiver and tropospheric delay, i.e.,

$$N_W = N_1 - N_2 = \varphi_1 - \varphi_2 - \frac{f_1 - f_2}{f_1 + f_2}\left(\frac{P_1}{\lambda_1} + \frac{P_2}{\lambda_2}\right) \tag{30.2}$$

and

$$\begin{aligned}\varphi_{LG} &= \lambda_1\varphi_1 - \lambda_2\varphi_2 \\ &= \left(\frac{I}{f_2^2} - \frac{I}{f_1^2}\right) + (\lambda_2 N_2 - \lambda_1 N_1)\end{aligned} \tag{30.3}$$

30.2.2 Multipath Effect Equation

For the wave length of carrier wave are much smaller than pseudorange, its multipath can be ignored when the multipath value of pseudorange is calculated.

By using 2 receivers in a same site in adjacent days, 2 observation files and navigation files are recorded. The site is called 0310 where a CORS station are planned to be built recently. With consideration of multipath and receiver noise, the Eq. (30.1) can be changed to the equation of multipath, i.e.,

$$\begin{aligned}MP1 &= P_1 - \varphi_1\lambda_1\tfrac{\alpha+1}{\alpha-1} + \varphi_2\lambda_2\tfrac{2}{\alpha-1} \\ MP2 &= P_2 - \varphi_1\lambda_1\tfrac{2\alpha}{\alpha-1} + \varphi_2\lambda_2\tfrac{\alpha+1}{\alpha-1}\end{aligned} \tag{30.4}$$

where $\alpha = f_1^2/f_2^2$, MP1 and MP2 are consist of the multipath of L1 and L2 carrier wave, receiver noise and the product of each wave length and ambiguity.

30.3 Calculation and Analysis

As expressed in Eqs. (30.2), (30.3) and (30.4), calculating the multipath value of each epoch needs cycle slips detection, bad epoch deletion, ambiguity elimination, etc. To analysis fully and accurately, and make the program compatible with different observation files, the flow chart is shown in Fig. 30.1. What' more, it also can read TEQC result, extract GPS information and convert into statistical table and ring view figure of site's multipath value.

30.3.1 Comparison with TEQC

As described in Fig. 30.1, same observation file and navigation file are used to compute multipath value both in TEQC and the designed program. TEQC which is a powerful pre-treatment software can output 9 files, including elevation, azimuth, multipath of L1 and L2 wave, and conclusion files. In this case, same files with universal format of above-mentioned files, are output from the software this paper designed. And the comparison of multipath of L1 and L2 wave between TEQC and

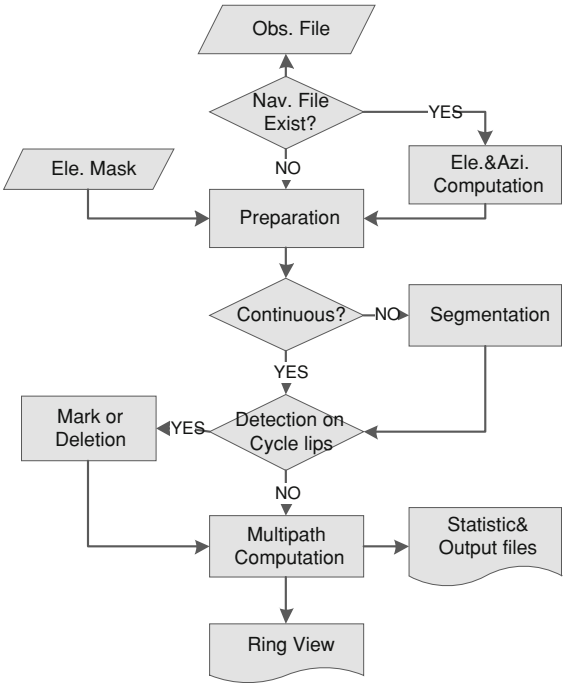


Fig. 30.1 Flow chart

this paper is drew in Fig. 30.2. The raw data are recorded in Site.0310, where is ready to build a permanent operational station and located in JiangXi, China.

Coincidences when compare multipath of each satellite occur in satellites which multipath value is relatively normal. And the biggest difference of calculation is less than 0.096 m, which lead about 0.009 m difference to the site' multipath RMS value.

For the location of site, the projection path of satellite orbit can be different. In other words, the multipath of few satellites is influenced for its lower elevation. Cycle slips or gross error easily occur when satellite run lower and its elevation are below a value decided by the environment of site. Different algorithms have distinct rules to detect these epochs. What this distinction result in is that continuous epoch are divided into different pieces. In the Fig. 30.2, the biggest difference occurs in the calculation of MP2 of satellite PRN.12. By comparing all epochs of PRN.12, and relating the elevation of each epoch, Fig. 30.3 concerning the different 2 algorithms is figured. There is little distinction between the TEQC result and this paper's result where are between 337 and 362 epoch, and can be saw in Fig. 30.3 red rectangle. After analysis on raw data, inappropriate detection are made by TEQC's algorithm, which case these epochs' calculation minus a wrong ambiguity and lead to a bias on multipath value.

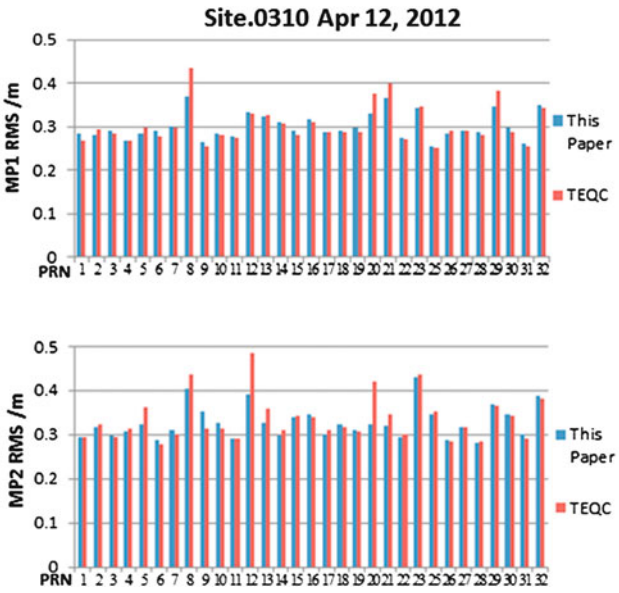


Fig. 30.2 MP1 and MP2 comparison

30.3.2 Contrast between 2 Antennas

With verification on the progress of calculating multipath value, the data recorded by 2 different receivers on 2 adjacent days are used to analysis the multipath effect on 2 different antennas. One data is recorded by Trimble R8 model 2 with TRM60158.00 antenna on April 12, 2012. Another is recorded by Trimble NetR5 with TRM55971.00 antenna on April 13, 2012. Both data are recorded on site 0310. By giving elevation mask such as 10, 15 and 20, calculation on different elevations can be drew as Fig. 30.4, and the statistical information are showed on Table 30.1.

In Fig. 30.4 and Table 30.1, multipath RMS of each satellite on 10, 15 and 20° elevation mask are be took into comparison. With consideration of the same site and the record data are two adjacent days, the environment can be consider as the same, and condition of the obstacle are same. On this premise, results can be described as follows:

1. Multipath value is relevant with elevation of satellite even if the data are recorded by different receivers, and it's smaller when elevation gets higher.
2. With the observation on the same site, the difference of multipath effect on these 2 antennas is small, and less than 0.1 m.
3. With the advantage of designation on antenna, Trimble NetR5 can reduce the multipath effect in lower elevation condition. Especially with the rising of elevation, the observations are improved better than the Trimble R8 model 2 GNSS system.

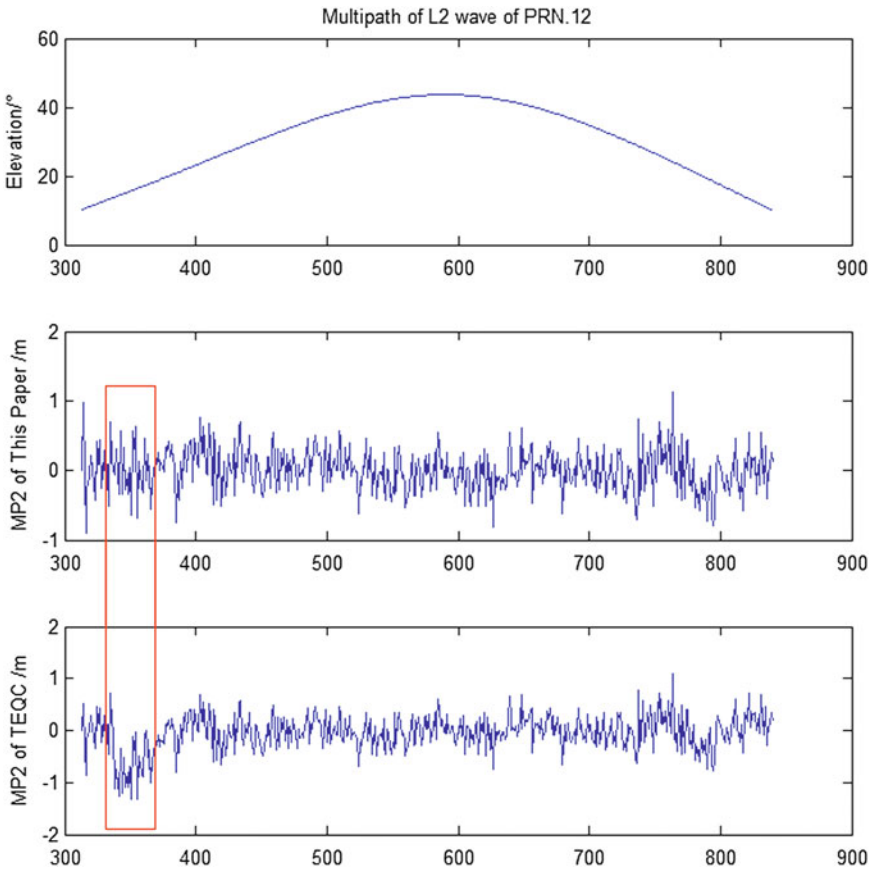


Fig. 30.3 Comparison on MP2 of PRN.12

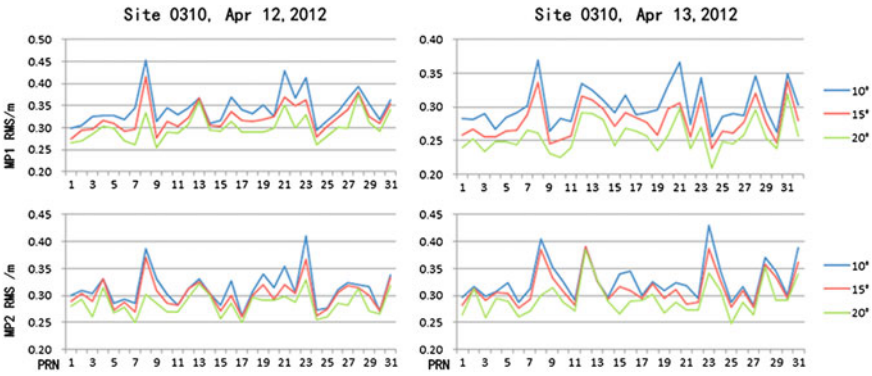


Fig. 30.4 Multipath value of trimble R8 model 2 and trimble NetR5

Table 30.1 MP1 RMS statistics

Receivers	Resources	10°	15°	20°	Adv. 15°/10° (%)	Adv. 20°/10° (%)
Trimble R8 model 2	This paper	0.35	0.323	0.3	7.71	14.29
	TEQC	0.348	0.322	0.298	7.47	14.37
Trimble NetR5	This paper	0.303	0.28	0.257	7.59	15.18
	TEQC	0.306	0.28	0.254	8.50	16.99
Advantages	This paper	13.43 %	13.31 %	14.33 %		
	TEQC	12.07 %	13.04 %	14.77 %		

Using other functions of this paper's software, qualities of observations are summed up as the Table 30.2, and the ring views of multipath of L1 wave are drew as Fig. 30.5.

What Table 30.2 shows verified the result that Trimble NetR5 has a better quality in observing. No matter in 10 or 20° mask of elevation, Trimble NetR5 has presented reliable characters that the data recorded by this receiver have less cycle slips and gross errors while recording more effective epoch. Especially in lower elevation in Fig. 30.5, the red points which mean multipath value is bigger than 1 m, are much more in data recorded by Trimble R8 model 2.

And conclusions can be drawn on this test. In the 10° mask of elevation, the reflections form east and south are the major contributor of multipath effect on site 0310. Due to first-class designation and the outstanding quality of observation, Trimble NetR5 receiver and TRM55971.00 antenna have a much better performance on reducing multipath effect than the contrast. For that observations from 2 receivers and antennas are very close in higher elevation, multipath effect can be considered as interference error which mostly occurs in lower elevation in positioning.

As a result, to get higher quality observation on this site where a permanent station will be built, receivers with satisfied performance such as Trimble NetR5 should be employed. Moreover, obstacle from east and south need appropriate disposition. All can be effective in the improvement of CORS service.

Table 30.2 Qualities of trimble R8 model 2 and trimble NetR5

Receivers	Ele. mask	Obs.	Bad	Slips	Gross	Total time
Trimble R8 model 2	10	24693	137	11	57	204:09:30
	15	22419	34	7	29	186:18:00
	20	19749	0	5	16	164:26:30
Trimble NetR5	10	24704	154	3	28	204:21:00
	15	22426	49	3	9	186:24:00
	20	19748	0	3	3	164:32:30

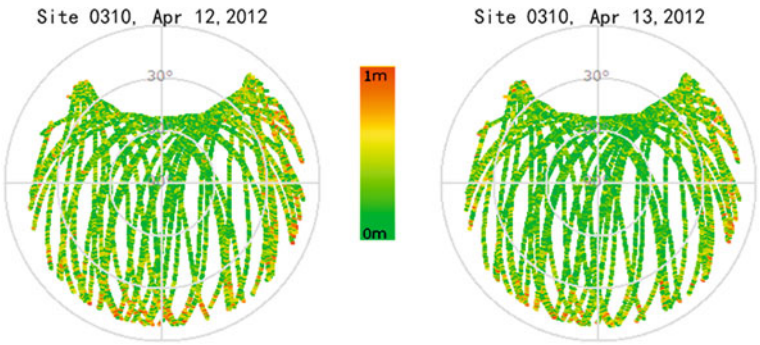


Fig. 30.5 Ring views

30.3.3 Multipath Effect Stability Analysis

With the result in Sect. 30.3.2, Trimble NetR5 is better than Trimble R8 model 2 when dealing with multipath effect. However, much more calculations are needed to prove the steadiness of Trimble NetR5. To do this experiment, data from an acknowledged site are necessary. ASPA is an IGS station, located in the Pacific Ocean, which is employed with Trimble NetR5 and TRM55971.00 antenna.

A month data of ASPA are downloaded from IGS website dating from February 1st–28th, 2012, and the multipath effect of each day are computed. By deducting the average of month, the result is showed in Fig. 30.6. What needs to be paid attention is that the data on February 3rd is dating from 0 to 22 o'clock and the reason is unknown.

In Fig. 30.6, the average is 0.6527 m, and the sigma is 6.72 mm. The difference with average of each day is less than double sigma. It can be considered that the result is in keeping with random pattern. The key reason leading to these little differences are from 2 sides. One is the receiver noise, and another is the influence caused by the accuracy of pseudorange. In summary, Trimble NetR5 has achieved a reliable performance in dealing with multipath effect, and can be employed in permanent operational station satisfactorily.

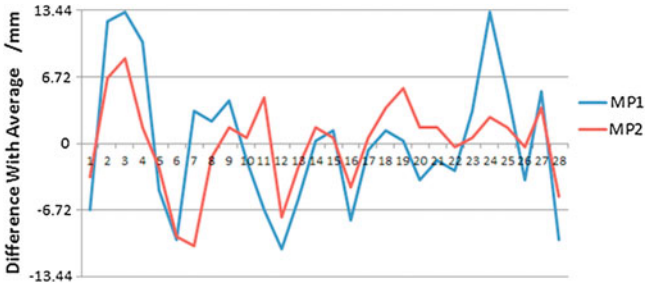


Fig. 30.6 Multipath time series on ASPA

30.4 Concluding Remarks

Multipath effect is an error of reflection in GNSS positioning. It is decided by the environment. However, the calculation on multipath have differences between 2 receivers and antennas this paper mentioned. In other words, with first-class designation and state-of-art technique of antenna, multipath effect can be reduced efficiently.

To tell the distinction between two commonly used Trimble receivers, a calculation on multipath effect was done. There are three major steps in implementing multipath analysis. The first step is to verify the practicability of the software designed in this paper, and the contrast with the result calculated by TEQC is finished. In the second step result of data recorded by two different receivers is compared. So is the statistic on the quality of observation. In the last step, the multipath time series are computed. With a month result, the better receiver, Trimble NetR5, in step 2 is given an estimate that its performance on multipath effect can be described as stable and reliable. Such a strategy is easily applied to others sites and receivers, and the software can be easily used in analysis on other data. Analysis on multipath effect of these two receivers and antennas, in some extent, can resolve questions on project. However, research on multipath effect still need extended.

References

1. Rongxin F, Chuang S, Wei N, Zhao QL (2009) Research on detection of real-time cycle slips in quality control of GPS data. Geometrics and information science of Wuhan University, Wuhan, Sept, vol 34, No. 9
2. Xia L (2001) Theoretical research and numerical results regarding multipath in GPS observation. Wuhan University, Wuhan
3. Ge L, Han S, Rizos C (2000) Multipath mitigation of continuous GPS measurements using an adaptive filter. GPS Solutions 4(2):19–30
4. Blewitt G (1990) An automatic editing algorithm for GPS data. Geophys Res Lett 17(3): 199–202

Chapter 31
The Study on Movement Characteristics and Non-linear Model of CGCS2000 Framework

Xiaoming Wang, Yingyan Cheng, Zhihao Jiang, Fuli Wang, Xing Chen and Xiaochao Li

Abstract It is essential to establish Non-linear model of CGCS2000 framework to improve the precision of domestic coordinate system, which can provide support for precise orbit determination of Beidou satellites and precise point positioning. And it can also meet the needs of high precision research and real time engineering project. Much research has been done about the movement characteristics and noise characteristics of CORS time series, however, this study only focus on how to improve the precision of linear velocity estimation on the colored noise properties of CORS Network and this method used has some flaws and limitations which may lead to unreliable result. In this paper, Singular Spectrum Analysis is introduced and based on which new interpolation and method, gross error detection method and signal–noise separation method has been introduced or proposed. The time series of domestic CORS network in china has been analyzed, as well as some IGS stations affected by earthquake. In this study, the SSA-P method is also proposed to produce more accurate forecast than traditional methods in current use.

Keywords CGCS2000 · Non-linear model · SSA · Time series

X. Wang (✉) · Y. Cheng · F. Wang · X. Chen
Chinese Academy of Surveying and Mapping, Beijing 100830, China
e-mail: chwxm@sina.cn

Z. Jiang
National Geomantics Center of China, Wuhan University, Beijing 100048, China

X. Li
Qingdao Institute of Geotechnical Investigation and Surveying Research, Qingdao 266032, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_31,

31.1 Introduction

A Terrestrial Reference Frame (TRF) is a set of physical points with precisely determined coordinates in a specific coordinate system (cartesian, geographic, mapping…) attached to a Terrestrial Reference System. Such a TRF is said to be a realization of the TRS. To do research on movement characteristics and model of framework using the latest mathematical and surveying techniques to attempt to realize the TRS as precisely as possible is very important. It will improve the precision of domestic coordinate system by establishing non-linear model of CGCS2000 [1–6] coordinate frame, which include two aspects,the first aspect is studying the deformation modeling under the loading of surface water/atmosphere mass, the second aspect is studying the non-linear movement characteristics of stations which is affected by earthquake, which is always difficult to be modeled with using short, noisy time series. Some research has been done to study the method of linear velocity estimation of CORS Network in China based on the CGCS2000 frame. There are also many studies about non-linear movement characteristics of CORS network in china based on traditional methods. However, few studies on non-linear model of CGCS2000 Framework have been published.

Singular Spectrum Analysis (SSA) is introduced in this paper, which is based on principal component analysis in the vector space of delay coordinates for a time series. This method has been used in oceanography for years. In this study, singular spectrum analysis for missing data (SSA-M) is introduced and Singular Spectrum Analysis-Inter Quartile (SSA-IQR), a gross error detection method, is proposed. These two methods are used for pretreatment of GPS time series. After that, SSA is combined with nonparametric test of Kendall and the maximum entropy method (MEM) to refine the interpretation of trend and oscillatory behavior. At last a prediction method, singular spectrum analysis for prediction (SSA-P), is proposed, which is used in this paper to establish forecast model for CGCS2000.

31.2 Singular Spectrum Analyses

For a time series $x_i$where index i varies from 1 to N, we choose an embedding space of dimension M ($0 < M < N$), in which a Toeplitz lagged correlation matrix can be constructed. Then the eigenvalues and eigenvectors of this matrix are determined and sorted in descending order of eigenvalues. The principal components (PCs) are defined as the orthogonal projection coefficient of the original series onto the eigenvectors of Toeplitz lagged correlation matrix which can be calculated with using method Classical PCA. However, PCs are processes of length N–M + 1 which cannot allow an unique expansion of the series into a sum of the different components. Therefore, we calculate reconstructed component (RC) which is of length N and can localize the short spells of the oscillations in the signal. By analyzing the

characteristics of RCs, we can determine a set of RCs to reconstruct the signal of original time series. More details can be found in paper [7].

31.3 Data Processing

In this research, we use Gamit/Globk software to process the data of 29 stations in the Crustal Movement Observation Network of China (CMONOC) to obtain the time series of this stations based on CGCS2000. Firstly, the data of 29 fiducial stations and 16 IGS stations around China was processed with Gamit. Thus we could get the h-files which contain estimates and an associated covariance matrix of station positions and orbital and Earth-rotation parameters. Secondly, we combine these h-files with SOPAC global h-files that were loosely constrained solution of stations from IGS network, to produce new solutions. Thirdly, the globk is run to process the data from combined h-files using method of free-adjustment with parameters loosely constrained. Finally, the glorg would be run to make the reference of the solution specified by using 92 reference stations. There were 92 reference stations selected to define the reference frame. Firstly, the quality of the data at each station was analyzed according to six principles including continuity principle, stability principle, high precision principle, muti-solution principle, balance principle and precision coherence principle. Then the stations were rigorous selected using the method of stations 7 parameters transformation. The third step is supervision clustering which can eliminate the stations that were largely affected by local diastrophism. Finally, we make the special distribution of stations well proportioned. The plane coordinate component repeatability and the height coordinate component repeatability of the daily result that were computed in the geodetic reference frame, ITRF2005 (International Terrestrial Reference Frame), respectively reaches 5 and 12 mm.

31.4 Analysis of Characteristics of Domestic CORS and IGS Stations

The observation data of CORS stations over several years reflect trend and oscillations and white noise and red noise are the basic characteristics in CORS stations [8]. If we want to get reliable result of this characteristic, the most important thing is to accomplish signal–noise separation correctly. According to traditional methods, the trend is assumed as linear and the main oscillations can be detected by analyzing the time series using power spectrum method. Then the signal–noise separation is conducted with using the formula (31.1), where a is constant, b is linear velocity, c, d, e and f are the parameters of oscillation, and $\sum_{j=1}^{n_g} g_j H\left(t_i - T_{gj}\right)$ is used to eliminate the effect caused by earthquake.

$$y(t_i) = a + bt_i + c\sin(2\pi t_i) + d\cos(2\pi P t_i) + e\sin(4\pi P t_i) + f\cos(4\pi P t_i) + \sum_{j=1}^{n_g} g_j H\left(t_i - T_{gj}\right) + v_i \tag{31.1}$$

However, this methods have some limitations which are detailed as follows: firstly, the previous study has proved that the noise in CORS are charactered as having white noise and red noise, so the result obtained based on white noise assumption is unreliable. Secondly, the trend, which is affected by many factors, is always non-linear. Thirdly, apart from periodic components in year and semi-year, there are also some high frequency oscillations that can be detected from signal; Finally, according to traditional method, the amplitude of oscillations is assumed to be a constant which means that the description of characteristics of oscillations may be not accurate. Therefore, the all of those demonstrated the inaccuracy of the analysis with using traditional methods. In contrast, SSA can extract as much reliable information as possible from time series without using prior knowledge and assumption.

31.4.1 Data Preprocessing of Domestic CORS Network

The key operation of data preprocessing is gross error detection and interpolation. Due to the limitations of the tradition methods to accomplish signal–noise separation, it may be unreliable to detect the gross error with analyzing the noise obtained using formula (31.1). And the accuracy of interpolation result using traditional methods depends on the property of missing data. Apart from that, the process is also complicated and may introduce personal error.

As traditional methods have some flaws and limitations, the SSA-IQR is proposed to detect gross error and SSA–M [9] is used to fill the gaps of several data sets in this paper. SSA is a data-adaptive, nonparametric method, which can extract as much reliable information as possible without using prior knowledge about underlying physics of the system, based on the information it also provide the gross error detection and data interpolation models. In this paper, the experiment is performed to identify the reliability of the methods, and the incomplete data of CORS in china is reconstructed. To obtain high precision result, the inner- and outer-loop iterations are repeated using the cross-validated optimal parameters. These methods have proved to be of high precision and high automation. In addition, they also can eliminate personal error that may be introduced using traditional method. Even the time series has been affected by earthquake, we can reconstruct missing data and detect gross error using these two methods correctly. More details can be found in the following papers.

Table 31.1 shows the method to simulate the missing data with missing rate of 18–20 %. From Table 31.2, the interpolation of 4 CORS, we can find that even the missing rate is so high and the length of missing data is more than 7 months, the

Table 31.1 Data simulation method

Num	Method	Start yr day	End yr day	Length (day)	Constant missing
1	Constant	2002 47	2002 147	100	100
2	Constant	2004 167	2005 1	200	200
3	Constant	2005 301	2005 331	30	30
4	Constant	2006 306	2006 336	30	30
5	Random	2002 192	2003 232	84	15

Table 31.2 Result of interpretation for CORS data

Site	Direction	Average error (mm)	Standard error (mm)	Site	Direction	Average error (mm)	Standard error (mm)
LHAS	N	2.0	2.5	WUHN	N	1.9	2.6
	E	2.2	3.7		E	1.8	4.3
	U	5.2	7.0		U	7.1	8.8
SHAO	N	1.7	2.1	BJFS	N	2.5	3.2
	E	2.0	2.6		E	3.8	4.7
	U	5.6	7.2		U	5.6	7.1

accuracy of interpolation in the horizontal direction is better than 5 mm, mostly superior to 1 cm in the elevation direction. Considering the noise level of the site itself, this method can be recognized as has high accuracy and stability.

Figure 31.1 shows the result of interpolation and gross error detection of BJFS station with the data observed from 1999 to 2009. The green part are residuals of interpolation using SSA-M and the red circles are gross error detected with SSA-IQR. SSA-M can obtain the signal model based on SSA, then the missing data can be reconstructed with the corresponding data of the model. Therefore, we can get the model residuals throughout the entire period of time series, which can be used as a criterion to analysis the precision of interpolation. In fact, this is one of advantages of this method to be used in practical interpolation process. To prove new methods is better than traditional methods, in this study, two traditional methods, orthogonal polynomial interpolation method and cubic spline interpolation method, are also be studied. Take BJFS station as example, the result shows that the accuracy will decrease to 2 cm and the accuracy at some epoch even reach 4 cm when the length of gap is over 2 months using traditional methods. Moreover, the interpolation result of some other stations will be distorted when the length of gap over 15 days.

31.4.2 Signal–Noise Separation and Analysis of Movement Characteristics of Domestic CORS

Accomplishing signal–noise separation of CORS time series correctly is crucial to obtain reliable analysis result of dynamic movement characteristics and noise

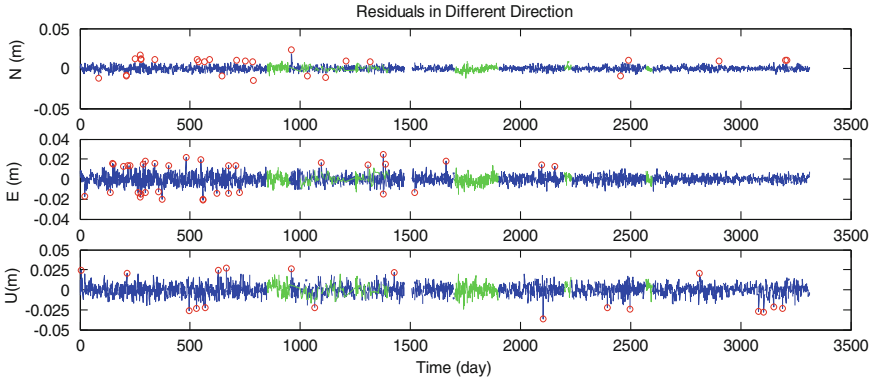


Fig. 31.1 Interpolation diagram of residual series at BJFS

characteristics. In this paper, the methods based on SSA are used to achieve this. The methods can be detailed as following: Firstly, the RCs are obtained by decomposing preprocessed time series using SSA, and then nonparametric test of Kendall is used to identify the trend item which will be removed from time series. This procedure will be repeated until no trend item can be identified from time series. Secondly, the time series without trend items is decomposed using SSA, then RCs are analyzed to identify oscillations. According to the principle of SSA, when a vigorous-albeit irregular oscillation present, then, (1) the two adjacent RCs must be localized around the same frequency; (2) the sum of the power spectrum energy between them is big enough. These two criteria are applied to the realizations of time series of CORS. Thirdly, the RCs identified as are analyzed using MEM. Finally, the RCs which are recognized as trend item and oscillations are used to reconstruct the signal.

By analyzing the data of CORS using this method, we can find that there are periodic components in year and semi-year in the data of most stations, and obvious periodic components in three months can also be detected in some stations. Taking result of BJFS station in vertical direction as example, the time series is shown in Fig. 31.2, in which the red part represents the trend item recognized using SSA method. From Fig. 31.2, it can be seen that the trend item is obvious nonlinear. Figure 31.2 shows the coordinate component repeatability of BJFS station in vertical direction which is obtained by removing the trend item from original series. In this figure, the red part represents the oscillations identified using SSA. From Fig. 31.2, it can be seen that the amplitude of the periodic item changes over time. That is to say, the traditional method has obvious flaws in modeling oscillations by using sine and cosine. The residuals obtained by removing the trend items and oscillations from original time series are shown in Fig. 31.2. It can be calculated that the standard error is 3.48 mm and average error is 4.34 mm.

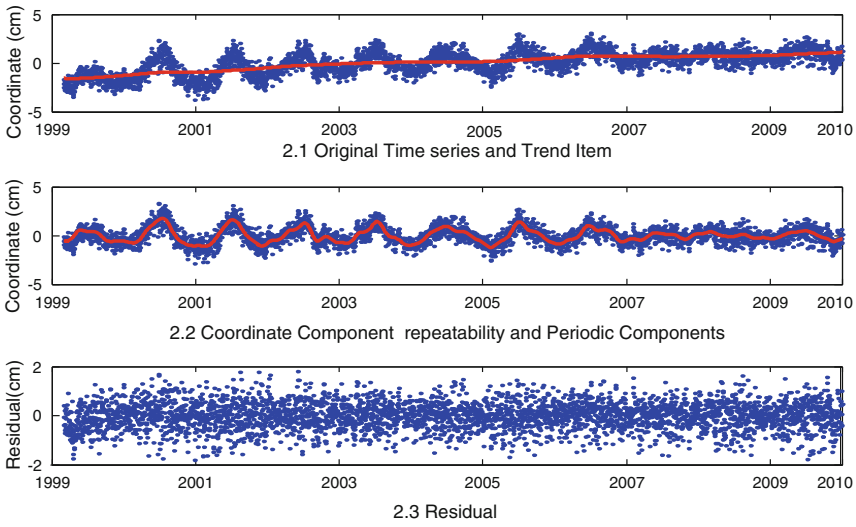


Fig. 31.2 The result of analysis at BJFS

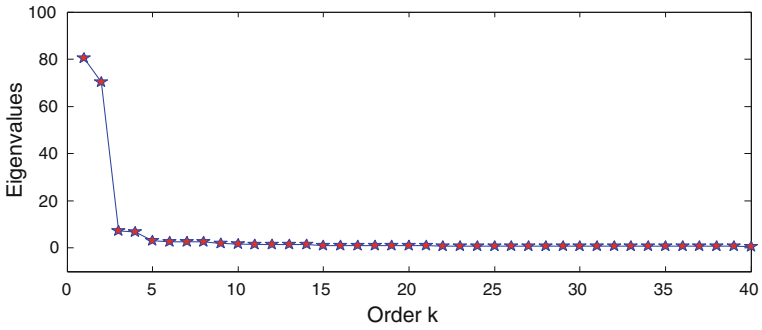


Fig. 31.3 The SSA eigenvalue spectrum of BJFS

Part of SSA eigenvalue spectra is plotted in Fig. 31.3. Figure 31.4 shows the results of the first nine RCs. According to principal of SSA, the greater the singular spectrum value is, the greater the proportion of signal. In Fig. 31.3, singular values of RC2 and RC3 come in pairs; therefore, they can be identified as oscillation by step 3. From Fig. 31.5, We can find that the sum of two RCs can be identified as annual cycle, that is, the superposition of RC2 and RC3 can reconstruct to the original annual term of the raw time series. The result of the power spectrum (Fig. 31.5) indicates that the site still has semiannual and seasonal oscillations. Using this method, the cumulative variance rates of the trend and annual terms are 40 and 35 % respectively, and he cumulative variance rate of the signal reconstructed with trend item and oscillation is 80 %.

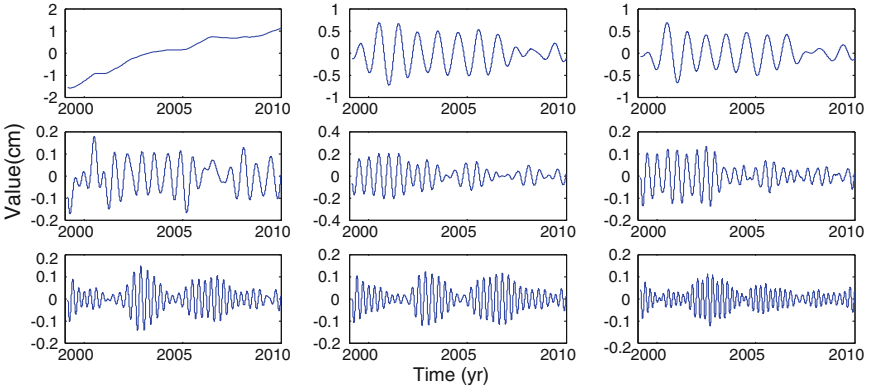


Fig. 31.4 Reconstructed Components of BJFS station

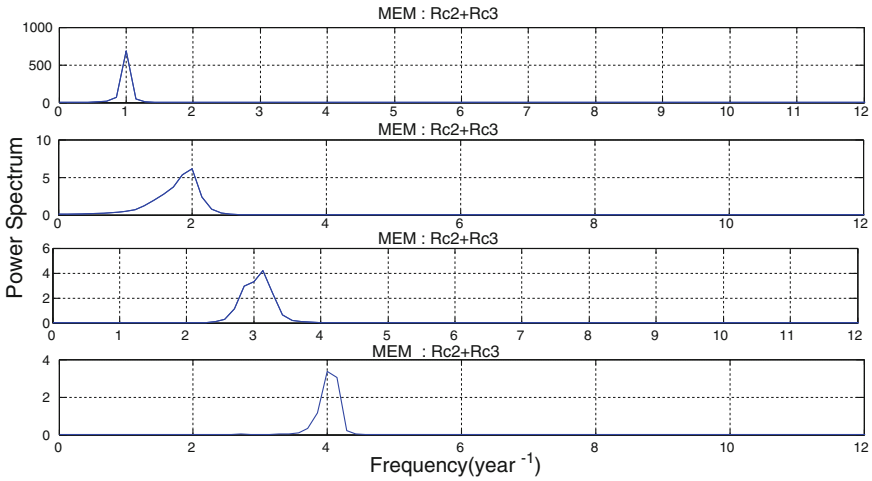


Fig. 31.5 Power spectrum used MEM

31.4.3 Analysis the Characteristics of the Domestic and IGS Sites Affected by the Earthquake

Sites' location could be teleported by earthquake through the instant energy release, and its velocity would be also changed after the earthquake. Figure 31.6 shows the time series of QION site affected by the M8.7 Indonesia Earthquake in 2004 (the moment black line corresponding to). Its liner velocity computed by GLOBK decreases by 2.3 mm/a after the Earthquake. The red component in this figure is the model obtained with using the method SSA. The fitting results show that the method is able to complete the high-precision modeling and possesses the

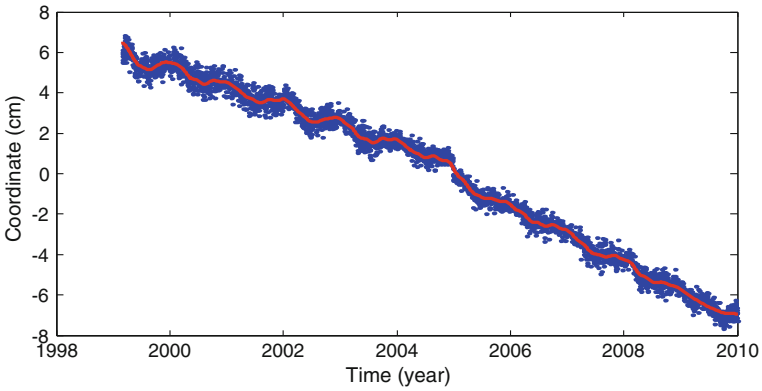


Fig. 31.6 The result of analysis at QION of the north

advantages of great adaptability even if there are mutant structural changes caused by earthquake in the time series. The analysis results of some domestic CORS stations affected by the earthquake are shown in Table 31.3. The modeling method SSA has a very high accuracy with mean standard error in horizontal is 1–3 mm and in vertical is about 2–5 mm.

The analysis results of the time series of TSKB site after the 3.11 earthquake in 2011 using SSA are shown in Fig. 31.7. The standard error in horizontal is about 1 mm and that is 6.7 mm in vertical. The displacement in E direction reaches 8 cm during the first 45 days after earthquake and the site moved in the form of a very obvious index model in the early post-earthquake. In terms of velocity, it is significantly faster than that before the earthquake. Although it slows down two months later, the annual average velocity is still reach to 22 cm/a. Therefore, the influence on TSKB site of the earthquake is not only result in greater coseismic displacement (0.59 m in the east), but also accelerate the movement of the post-earthquake plate and also cause obvious nonlinear movement. The result shows that SSA can model the form of such a complex movement with high accuracy, reliability and self-adaptability.

Table 31.3 Result of analysis for CORS data

Site	Direction	Annual cycle	Half-year cycle	Standard error (mm)	Cumulative variance (%)	Earthquake
WUHN	N	Yes	Yes	1.0	99	MW8.0 Wenchuan earthquake
	E	Yes	Yes	1.1	99.4	
	U	Yes	Yes	2.6	63	
KUNM	N	Yes	No	1.7	99	Mw 8.7 Nias earthquake
	E	Yes	No	3.4	99	
	U	Yes	No	5.7	60	
DLHA	N	Yes	Yes	1.1	80.3	Mw 8.1 Kunlunshan earthquake
	E	Yes	Yes	1.4	99.4	
	U	Yes	Yes	2.5	65	

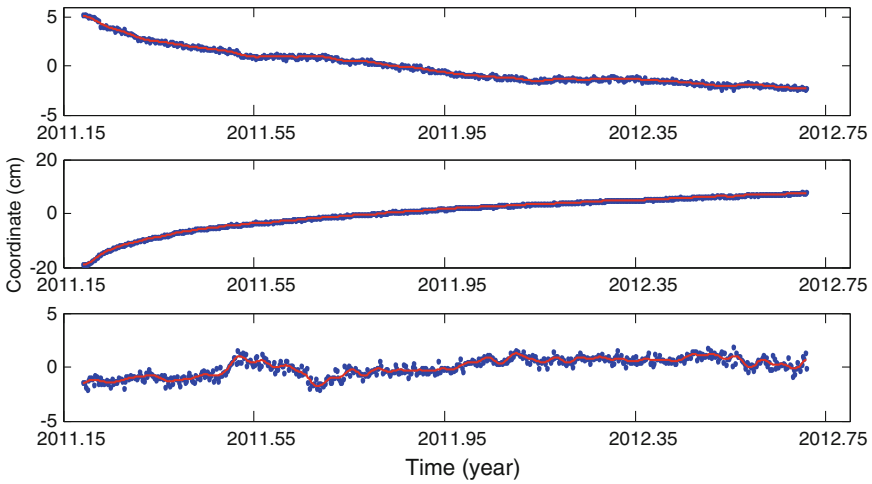


Fig. 31.7 The result of analysis at TSKB

31.5 Constructing Non-Linear Velocity Models Using SSA-P

Because the main function of SSA is to obtain data structure, this method is considered as cannot be used for forecast. Therefore, we always use the other models to produce forecast of individual PCs, which will be used with reconstruction algorithm to compare forecasts with real data. Essentially, any forecast method must be conducted based on obtaining data structure, so SSA method could be used to predict the data with acquiring the structural properties of the time series.

The basic ideas of Singular Spectrum Analysis for Prediction (SSA-P) are described as follows: First of all, analyzing the time series using SSA, determine the order P which presents the number of RCs used to reconstruct the signal and complete the signal reconstruction. Next, adding data of length of K to the end of the signal model reconstructed with P RCs. The initial value of added data can be determined by linear model constructed with the end of signal. Then, the new series is analyzed using SSA, and the values of added data are replaced with the corresponding values in the RC1. It is an iterative process until convergence. Fourthly, decompose the new sequence by SSA, take RC1 + RC2 as the new values to replace the k values at the end. It is also an iterative process until convergence. Repeating the above steps until P RCs are used to reconstruct K data which is the forecast data we want to obtain. The time series of URUM site before 2011 are used to build model (the green line) and predict the series for the next 1.5 years (the red line), and the results are plotted in Fig. 31.8. The mean square error is 3.16 mm and the average error is 4.6 mm.

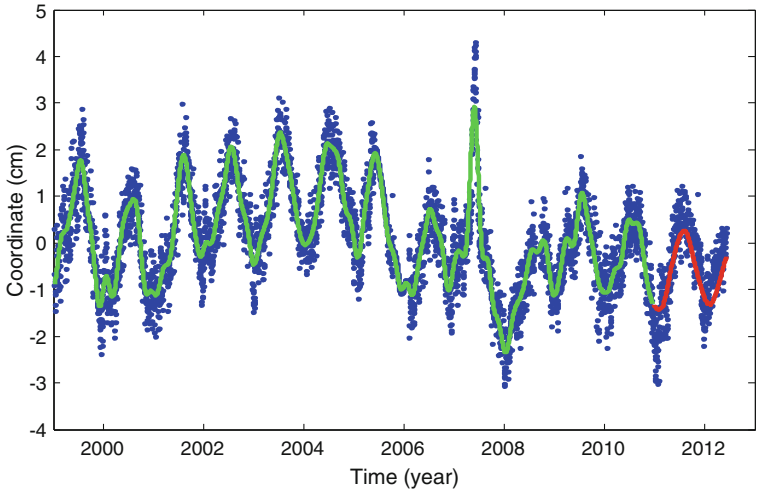


Fig. 31.8 The result of prediction at URUM stations of the vertical

31.6 Conclusions

1. Establishing Non-linear model of CGCS2000 framework can improve the precision of domestic coordinate system, which can provide support for precise orbit determination of Beidou satellites and precise point positioning. Based on studying the theory of SSA, we propose some new methods to resolve the questions that may present when we use SSA in practical experiment. For example, the pseudo-data method is proposed to resolve the phase-shift at the end of data when we reconstruct the signal. In this paper, we also propose SSA-IQR and SSA-P to accomplish gross error detection and prediction respectively. This new methods are combined with SSA-M and MEM methods to establish non-linear model of CGCS2000 framework. The new methodology has been proved to be of high precision, reliability and self- adaptability. More details can be found in following papers.
2. Because the kinetic characteristic of sites affected by earthquake relates to a variety of factors, modeling often has a certain complexity. The modeling method proposed in this paper has the advantage of great self-adaptive, which can ensure we can get a high precision model with analyzing short, noisy time series and reestablish the coordinate frame of the areas affected by earthquake quickly and accurately.
3. Because loads of research programs and construction project has been conducted, the number of CORS in china has increased dramatically these years. Therefore, it is very important to study the best model to ensure the data be used effectively. This is one of main aims of our study. And how to establish non-linear velocity field using so many CORS that will be constructed in future is another project we will study next.

Acknowledgments This research is funded by the National '863 Geocentric Coordinate Project' (No: 2013AA122501), the 'Promotion and Application of CGCS2000' (No: A1201).

References

1. Junyong C (2008) Chinese modern geodetic datum. Acta Geodat-ica et Cartographia Sinica 37(3):269–271
2. Cheng P, Wen H, Cheng Y, Wang H (2009) Parameters of the CGCS2000 ellipsoid and comparisons with GRS80 and WGS84. Acta Geodat-ica et Cartographia Sinica 38(3):189–194
3. Ziqing W (2008) China geodetic coordinate system 2000 and its comparison with WGS84. J Geodesy Geodyn 28(5):1–5
4. Liu J (2009) Some thoughts on the establishment of nationwide continuously operating reference stations. Geomatics Inf Sci Wuhan Univ (34):1261–1265
5. Yang YX (2009) Chinese geodetic coordinate system 2000. Chinese Sci Bull 54(15):2714–2721. doi:10.1007/s11434-009-.342-9
6. Cheng PF, Cheng YY, Wen HJ et al (2008) Practical manual on CGCS2000. Press Surveying Mapp, vol 10, Beijing
7. Vautard R (1992) Singular spectrum analysis: a toolkit for short, noisy chaotic signals. Physic D (58):95–126
8. Jiang Z, Zhang P et al (2010) Velocity estimation on the colored noise properties of CORS network in China based on the CGCS2000 frame. Acta Geodaetica et. Cartographica Sinica 39(4):355–362
9. Kondrashov D, Ghil M (2006) Spatio-temporal filling of missing points in geophysical data sets. Nonlinear Process Geophys 13(2):151–159

Chapter 32
Application of Curve Surface Fitting in Regional Ionospheric Delay Model with Sparse Station Distribution

Enqiang Dong, Weijie Sun, Xiaoli Wu, Jing LI and Jiachen Fan

Abstract Ionospheric delay observations of regional satellite navigation system shows heterogeneity in space and time, and it brings difficulties to implement the ionospheric delay model. The Kloubuchar model in geographic coordinates assumes that the ionospheric delay at the same latitude have the consistent variation characteristics with local time. The constellation of COMPASS is mainly composed of GEO satellites. In order to achieve the regional ionospheric delay variation, we can use several GEO ionospheric observations. By selecting 5 observation stations with well-distributed in latitude from 24 stations with COMPASS observations through the year 2012, we adopt polyhedral function fitting and Delaunay triangulation fitting to complement ionospheric observations of the region and period of time without real observations. Then we compute the Kloubuchar model parameters by using the complemented data, and use all 24 stations observations to check the results. Analysis result shows that, in the solar activity peak year, the average correction accuracy of model is prior to 65 % in middle-low latitude area, and RMS (1σ) is better than 11 TECU. The precision improvement is obviously improved compared with results without data complement using polyhedral function fitting.

Keywords Sparse station distribution · Ionospheric delay · Curve surface fitting · Polyhedral function · Triangulation

E. Dong (✉) · W. Sun · X. Wu · J. LI · J. Fan
Beijing Satellite Navigation Center, Beijing, China
e-mail: paradise4me@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_32,

32.1 Introduction

Ionospheric delay is an important error source of satellite navigation and positioning. The ranging error caused by ionospheric delay varies from 5 to 150 m, from the zenith to the horizontal direction [1]. For the reason that the ionospheric delay is a function of wave frequency, the dual-frequency users can eliminate ionospheric delay by dual-frequency navigation correction method, and single frequency users generally use the system broadcast ionospheric delay model parameter to modify the delay error. In the regional satellite navigation system, land limits of monitoring station distribution and faults of some satellites and stations will cause inhomogeneities of the ionospheric delay measurement in space and time. Scarcity of ionospheric observation leads to difficulties of ionospheric delay modeling implementation.

In Ref. [2], based on the daily changing characteristic of the ionospheric delay large scale changing related to the earth rotation, the author uses the methods of projecting latitude to local time to interpolate VTEC in a large scale area, so as to get comparatively perfect ionospheric delay space–time sample.

For the regional satellite navigation system, seriously insufficient of observation may cause divergence in the ionospheric delay model solution process, or precision significantly reduced. In the sun-fixed geographic coordinates, the Kloubuchar model equation assumes that the ionospheric delay of the same latitude have the consistent daily changing characteristic with local time, and the constellation of COMPASS is mainly composed of GEO satellites, so as to ensure 24 h continuous ionospheric observation. It brings turning point to solve the problem of insufficient measurement. Firstly, by projecting the IPP(ionospheric puncture point) geographical longitude to local time, we can obtain observations unrelated with the longitude but the local time; then, by using the curved surface fitting methods, we can obtain all or most of the ionospheric delay observation, so as to solve the divergence problem, at the same time, to improve the accuracy of the correction results of ionospheric delay model.

Using 2012 COMPASS observation, selecting 5 stations with well-distributed in latitude out of all 24 stations, we use curved surface fitting, such as polyhedral function and Delaunay triangulation, to complement data of the region without observations. Then we compute the Kloubuchar model parameters by the data after complement, and use the observations of all 24 stations to check the model results. Analysis results show that, in the solar activity peak year, the precision improvement is obviously compared with results without data complement.

32.2 Methods of Curved Surface Fitting

The familiar methods of curved surface fitting include polyhedral function fitting, Delaunay triangulation interpolation and spherical harmonic function fitting, etc. This paper adopts polyhedral function fitting and linear interpolation based on Delaunay triangulation to fit the regional ionospheric observations.

32.2.1 Polyhedral Function Fitting

Polyhedral function is raised by American professor Hardy in 1971, widely used in geodesy, gravity field, fitting Geoidal rise and crustal deformation. From the view of geometric point, using "either a regular or irregular surface can be made up of some simple surface with arbitrary precision to the superposition approximation" as the theoretical basis, it establishes specific function (called polyhedral function) in each interpolation point, with all the known data points respectively; then, these multi function value are added up to obtain the best surface fitting precision [3, 4]. Formula is:

$$\zeta = f(x, y) = \sum_{i=1}^{n} a_i Q(x, y, x_i, y_i) \tag{32.1}$$

In the formula, a_i are undetermined coefficients; x, y are undetermined point coordinates; x_i, y_i are the known point coordinates; $Q(x, y, x_i, y_i)$ are the second-order kernel functions of x, y. Generally, kernel function expression is:

$$Q(x, y, x_i, y_i) = [(x - x_i)^2 + (y - y_i)^2 + \delta]^k \tag{32.2}$$

The δ is smooth factor, and its value can be any non-negative; k is the exponential parameter, generally being chose as 0.5, 1.5 or −0.5.

32.2.2 Linear Interpolation Based on the Delaunay Triangulation

How to divide a scattered point set into irregular triangle mesh (Triangular Irregular Network, TIN) is triangulation problem of scattered point set. In all of the generated TIN method, Delaunay triangulation is optimal. It is as far as possible to avoid a morbid triangle, often being used to generate TIN [5], as shown in Fig. 32.1.

In the linear interpolation algorithm based on Delaunay triangulation, the input scattered point set is firstly converted to construct a Delaunay triangulation, and then new points will be inserted into the existing Delaunay triangulation network one by one so as to realize the interpolation arithmetic. For a given Delaunay triangular, its three vertex value are fi(xi, yi), where i = 1,2,3, and the new point (x, y) is located in the triangle. The equation defined by the three apexes of the Delaunay triangle plane is:

$$\mathrm{f(x,y)} = \mathrm{c1x} + \mathrm{c2y} + \mathrm{c3} \tag{32.3}$$

Linear equations can be obtained with the three known vertex fi(xi, yi), {i = 1,2,3}, to calculate the coefficients of [c1, c2, c3], and then the value of the new point (x, y) in the triangle.

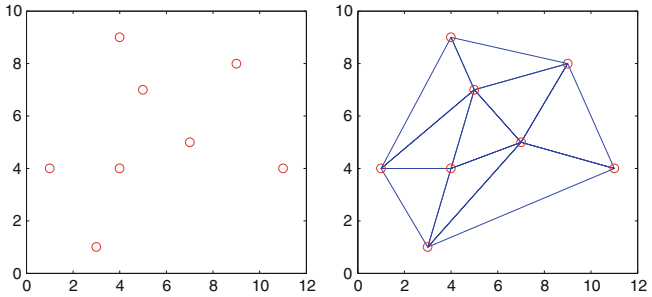


Fig. 32.1 Delaunay triangulation

32.3 Ionospheric Delay Model Realization

32.3.1 Observation Pretreatment

We select 5 stations observation well-distributed in latitude from 24 stations distributed in the territory to calculate the model parameters. They are Hulun Buir, Beijing, Hubei, Hainan and Sansha. Ionospheric delay is calculated by the methods of phase smoothing pseudorange, while the satellite and receiver BIAS are calculated with the methods of fixing a satellite TGD to its original calibration [6]. Using COMPASS VTEC data of the 298th day of 2012, the 24 h continuous IPPs map of the 5 stations is shown in Fig. 32.2.

Firstly, we project every IPP's longitude into local time in the sun-fixed geographic coordinates, that is:

$$Lt = t + 3,600 \cdot \lambda/15 \tag{32.4}$$

Among them, Lt (in second) is the local time, t (in second) is the observation time, λ (in degree) is the geographic longitude of IPP.

After projecting geographic longitude into local time, we make a grid density average compression with all IPP VTEC by 1° in latitude and 1 min in local time, that is to take the average value of all the IPP VTEC within the grid as the grid VTEC. If the grid is empty, its grid VTEC value is set to zero. The VTEC map of the 5 monitoring stations after compression is shown in Fig. 32.3a, while that of 24 stations is shown in Fig. 32.3b.

32.3.2 Curved Surface Fitting

We fit the 5 stations VTEC with polyhedral function curve fitting and linear interpolation based on Delaunay triangulation. Then, we compare the fitting results with all 24 stations VTEC results. Fitting error RMS of 6 days is shown in Table 32.1, and fitting results maps of the 298th day are respectively given in Fig. 32.4.

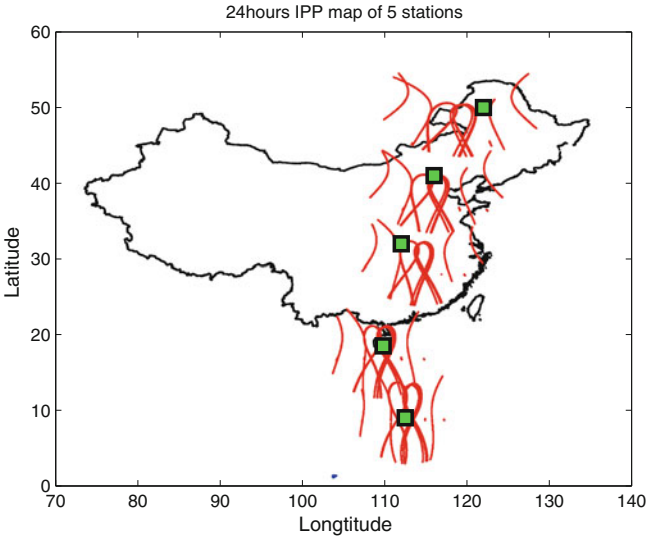


Fig. 32.2 24 hours IPP map of the 5 stations

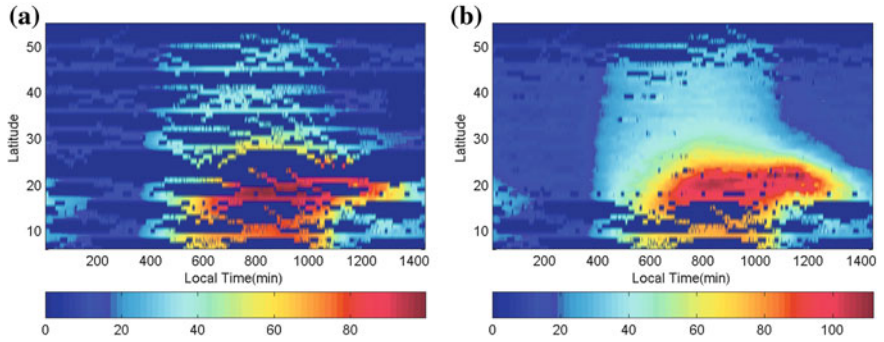


Fig. 32.3 VTEC after the compression. **a** 5 stations VTEC after compression. **b** 24 stations VTEC after compression

In polyhedral function curve fitting process, firstly, we compress the original data by grid of 1° in latitude and 3 min in local time. Secondly, we choose one point orderly every 13 IPPs as a kernel function points, and select 1/3 other IPPs as calculation points, taking 3 min as abscissa coordinate unit and 1° as the vertical coordinate unit. In polyhedral function fitting, selection of smooth factor is very important. In this experiment, we select the smoothing factor δ from 2 to 10, step 0.5, and calculate the fitting RMS one by one. When δ is admitted to 6.5, the kernel function matrix inverse used to calculate the kernel function parameters is singularity, and when δ is 6, RMS is minimum, 1.854 TECU.

Table 32.1 Curve surface fitting precision

Fitting methods	RMS(TECU)					
	271	272	273	298	299	300
Polyhedral function	1.882	1.729	1.977	1.854	1.792	2.135
Delaunay triangulation	1.721	1.596	1.801	1.682	1.618	1.789

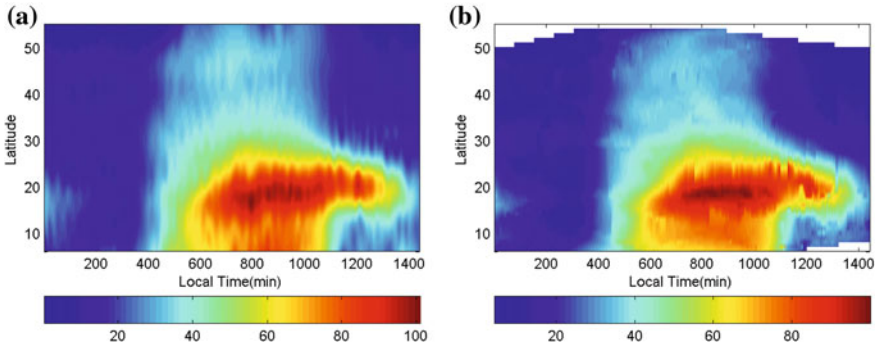


Fig. 32.4 Comparison of curve surface fitting results (The blank area in the second figure shows the disadvantage of Delaunay triangulation fitting. There will be no interpolation results out of the maximum envelope convex polygon). **a** VTEC after polyhedral function fitting. **b** VTEC after delaunay triangulation fitting

Linear interpolation based on Delaunay triangulation algorithm steps are as follows: (1) inputting scattered IPP(λ, φ) sets to construct Delaunay triangle net; (2) inputting the insertion Pi (λ, φ), finding the Delaunay triangle including Pi, then calculating the insertion point VTEC; (3) repeating step (2) until all the insertion points calculation are finished. The advantage of this method is that all known points are agonic, disadvantage is that there will be no interpolation results out of the interpolation area maximum envelope convex polygon. For this experiment, area beyond maximum envelope convex polygon rarely affects next calculation. After Delaunay triangulation linear interpolation, the VTEC error RMS is 1.682 TECU.

32.3.3 Ionospheric Delay Model Parameters Calculation

In the sun-fixed geographic coordinates, the characteristic parameters from VTEC daily curve can be extracted as A(cosine amplitude) and P(cosine cycle) [2]. The nighttime ionospheric delay is set to constant. The expression Kloubuchar is shown as [7]:

$$I_Z(t) = DC0 + A\cos\frac{2\pi(t - 50,400)}{P} \tag{32.5}$$

In the expression, $I_Z(t)$ is the ionospheric vertical zenith delay in seconds, and for the calculation of different frequencies, it needs to be multiplied with a frequency dependent mapping function K(f); t is the local time of the IPP, ans its unit is seconds; DC0 is the constant to express the nighttime ionospheric delay level.

A is the ionospheric delay cosine curve magnitude, in seconds, using alpha coefficients to be calculated;

$$A = \sum_{n=0}^{3} \alpha_n \phi_M^n \tag{32.6}$$

P is the period of cosine curve, in seconds, using beta coefficients to be obtained

$$P = \sum_{n=0}^{3} \beta_n \phi_M^n \tag{32.7}$$

φ_M is the geographic latitude of the IPP, and its unit is half of the week (π).

Using correlation processing method, we make correlation calculation between actual VTEC sequences and VTEC sequences calculated with 2 characteristic parameters on the same latitude, and calculate the correlation coefficient and RMS value. We firstly set 2 parameters ranges and calculation step value, and then we calculate some group correlation coefficient and RMS value step by step. When the correlation coefficient is maximum and RMS is relatively small, we confirm the group of 2 characteristic parameters of the latitude is the best.

In the course of calculating the final 8 ionospheric model parameters with VTEC observation, first of all, we transform every IPP's longitude into local time and make a grid density average compression with all IPP VTEC by 1° plus 1 min; so as to get VTEC daily variation time sequences of every latitude. And then, by using the correlation method, we calculate the 2 characteristic parameters sequences of every latitude belt; finally, in accordance with Kloubuchar ionospheric model formula, we establish a sequence of observation equation with the characteristic parameters sequences of every latitude belt and solved all 8 model parameters by least square method.

32.4 Results Analysis

In this experiment, 6 days observations have been analyzed, and they are the 271st, 272nd, 273rd, 298th, 299th and 300th day of 2012. We separately use 5 selected monitoring station observation data, all monitoring station observation data, linear interpolation results based on Delaunay triangulation and polyhedral function fitting results to solve the ionospheric delay model computation. Then, we

Table 32.2 Compare of all kinds of ionosphere model results

Model type	Doy = 271 (available points: 51,553)			Doy = 298 (available points: 51,316)		
	Err > 20 TECU	RMS (TECU)	correct percentage	Err > 20 TECU	RMS (TECU)	correct percentage
5 station	6,088	15.33	57.53	6,219	15.08	56.89
Triangulation	3,761	10.20	68.51	3,812	10.89	67.98
Polyhedral function	3,731	10.16	68.62	3,798	10.86	67.79
All station	5,078	12.77	67.36	5,106	12.38	66.73

undertake compare to check the results with observation of stations which did not participate in the model calculation, and statistical RMS of correction error and the average correct percentage. Because 2012 is the peak year of solar activity, we add statistics of points whose absolute correction error are greater than 20 TECU. All 6 days results are close, and Table 32.2 only lists the results the 271st and 298th day of 2012.

The formula to calculate the average correct percentage is:

$$Percentage = \frac{\sum_{i=1}^{n}\left(1 - \frac{abs(MVtec_i - Vtec_i)}{Vtec_i}\right)}{n} \tag{32.8}$$

Among them, n is total number of available observation, $Vtec_i$ is the original observation, $MVtec_i$ is the calculated value by model.

Analysis results show that, by using the data complemented by linear interpolation based on Delaunay triangulation and polyhedral function fitting, the model results are obviously improved compared with the results of 5 stations original observation, and a slight increase compared with the results of all monitoring stations observation. Results of two kinds of curve surface fitting are close, RMS is better than 11 TECU, the average correction percentage is prior to 65 %, and ratio with absolute error value greater than 20 TECU is about 7.5 %.

Figure 32.5 gives the 5 stations original VTEC after grid compression on the 20° north latitude (red), and VTEC after supplement with linear interpolation based on Delaunay triangulation (blue), and four kinds of model results of the same latitude are given together. In Fig. 32.5, it is obvious that all known points are unbiased in the linear interpolation results based on Delaunay triangulation, and two kinds of curved surface fitting model results are quite close, but better than that of the 5 station original model results, slightly better than all the station model results.

We also do interchanging analysis with other stations in the same latitude belt instead of the 5 monitoring stations, using Xinjiang station to replace Beijing station, Sichuan station and Tibet station to replace the Hubei Station, and the analysis results have no difference.

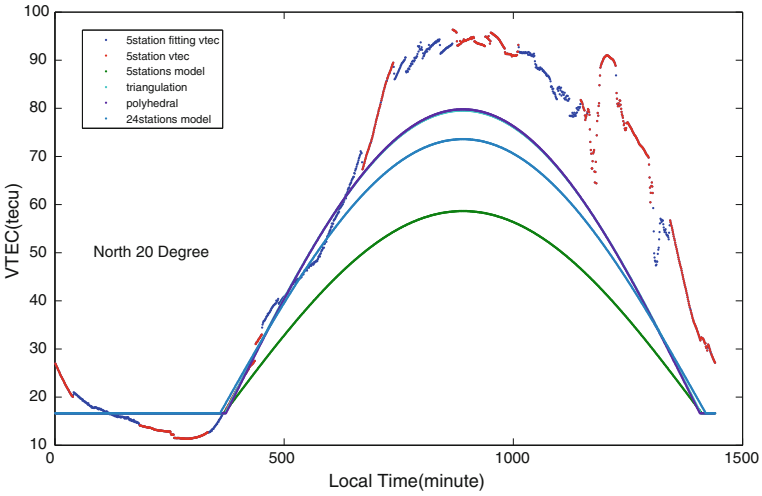


Fig. 32.5 North 20° model results

32.5 Conclusion

By selecting 5 observation stations with well-distributed in latitude from 24 stations with COMPASS observations through the year 2012, we adopt polyhedral function fitting and Delaunay triangulation fitting to complement ionospheric observations of the region and period of time without real observations. Then we compute the Kloubuchar model parameters by using the complemented data, and use all 24 stations observations to check the results. Analysis result shows that, in the solar activity peak year, the average correction accuracy of model is prior to 65 % in middle-low latitude area, and RMS (1σ) is better than 11 TECU. The precision improvement is obviously improved compared with results without data complement using polyhedral function fitting.

Linear interpolation algorithm based on Delaunay triangulation is simple, rapid, unbiased, and it is also applicable to wide area differential grid ionospheric region fit; polyhedral function fitting has been widely used in gravitational field, tidal model, geomagnetic field and many other fields, but in the application process, we need to constantly adjust the kernel function selection and determine the smoothing factor value, and further exploring application is also need to do in the region of ionospheric fitting.

References

1. Zhongmo Z, Jiejun YI (1997) Theory and application of GPS satellite measurement. Mapping Express, Beijing, p 01
2. Hongping Z (2006) Study of China regional ionosphere measurement and delay correction based on ground-based GPS. Doctor dissertation, Shanghai Astronomical Observatory; Chinese Academy of Sciences, 2006.05

3. Dajie L, Benzao T (2000) Process methods of practicality observation. Mapping Express, Beijing
4. Yifeng C, Hongzhou C, Min W (2012) Study on construction of local marine geomagnetic field model based on orthogonal least square multi-surface function. J Geodesy Geodyn, 2012.03
5. Lirong Y, Xingxiang J (2012) Delaunay triangulation interpolation algorithm and its application on MT mapping. Northwest Seismol J, 2012.03
6. Xiaoli W, Jinsong P, Liu L, Nan X (2011) Hardware delay solution of regional satellite navigation system. Geomatics Inf Sci Wuhan Univ, 2011.10
7. Klobuchar J (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aerosp Electron Syst AES-23:325–331. ISSN 0018-9251

Chapter 33
Precise Point Positioning Using GPS and Compass Observations

Wei Li, Peter Teunissen, Baocheng Zhang and Sandra Verhagen

Abstract The Compass Navigation Satellite system, which currently provides more than 12 satellites with three carrier signals, already satisfies the requirement of stand-alone positioning in the Asia–Pacific regional area. First an initial introduction and performance assessment of dual-frequency un-differenced precise point positioning (PPP) for GPS and Compass is presented, the results of which indicate that centimeter-level positioning accuracy of Compass-PPP is comparable to that of GPS-PPP. Then the combined GPS + Compass dual-frequency PPP model is introduced, followed by a numerical performance analysis and comparison with single GNSS-PPP. The results show that the combined GPS + Compass PPP can shorten the convergence time, but not necessarily improve positioning results by much if the satellites of the single GNSS system already have a good receiver-satellite geometry.

Keywords GPS · Compass · GPS + Compass, PPP · Dual and triple-frequency GNSS

W. Li (✉)
School of Geodesy and Geomatics, Wuhan University, Wuhan, China
e-mail: wewelee_c@whu.edu.cn

W. Li
Chinese Academy of Surveying and Mapping, Beijing, China

P. Teunissen · B. Zhang
GNSS Research Centre, Department of Spatial Sciences, Curtin University, Perth, Australia
e-mail: P.Teunissen@curtin.edu.au

S. Verhagen
Department of Remote Sensing and Geoscience, Delft University of Technology, Delft, The Netherlands

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_33,

33.1 Introduction

The Compass Navigation Satellite System is a global navigation satellite system, which is independently deployed and operated by China and still in development. Currently, Compass system consists of fifteen operational satellites transmitting navigation signals at three frequency bands (B1, B2, B3), including five operational Geostationary Orbit (GEO) satellites, five Inclined Geosynchronous Orbit (IGSO) satellites, and five Medium Earth Orbit (MEO) satellites. This enables Compass system to provide navigation service in the Asia–Pacific regional area; by 2020, a global navigation service will be achieved eventually. Worldwide users will be able to have access to services of Compass for positioning, navigation and timing (PNT), and also take advantage of multi-frequency observations from multi-GNSS systems to greatly enhance observation redundancy and the navigation performance [16].

With the increasing development of Compass system, it draws enormous interest and attention of the scientific community. Based on simulated data, Chen et al. [3] and Yang et al. [16] pointed out the contribution of Compass to user's PNT by analysis of visible satellites and dilution of precision (DOP) values. Absolute and relative positioning tests were also performed using simulated Compass observations ([1, 2]). Nadarajah et al. [10, 11] used multiple GNSS antennas mounted on a platform to determine the attitude precisely, with the constrained Least-squares Ambiguity Decorrelation Adjustment (C-LAMBDA) method. They demonstrated this method with real Compass data and quantify the improved availability, reliability, and accuracy of attitude determination using the combined GPS and Compass constellations. Since the Compass navigation message has not yet been publicly released, some community have estimated the orbit and clock of Compass GEO and IGSO satellites based on a network of Compass-capable tracking stations. For example, Steigenberger et al. [15] utilized the dual-frequency GPS (L1 and L2) and Compass (B1 and B2) data from the IGS Multi-GNSS Experiment (MGEX) and the Cooperative Network for GIOVE Observation (CONGO) to estimate the orbit and clock errors of Compass GEO and IGSO satellites, with the orbit quality at meter level for GEO satellites and one to two decimeter level for IGSO satellites. Shi et al. [13] from Wuhan University determined the precise orbit of Compass satellites, with the radial accuracy better than 10 cm. These two communities also pointed out that higher accuracy orbit and clock of Compass satellites can be achieved by denser distributed tracking stations.

Along with the accomplishing of regional Compass system, investigations have focused on positioning performance using real observations mostly. Shi et al. [13, 14] showed that an accuracy of 20 m can be achieved for Compass standalone positioning; in short baseline experiment, the precision of Compass-only relative positioning are 2–4 cm; and the combined GPS + Compass solutions have improved the positioning by at least 20 % compared with GPS-only solutions. Based on achieved orbit and clock products, static PPP and kinematic RTK can also achieve centimeter level and 5–10 cm respectively. Steigenberger et al. [15]

also tested the Compass PPP performance and compared it with GPS PPP in the aspect of positioning and zenith wet delays (ZWD) solutions. Montenbruck et al. [8, 9] demonstrated the high level of stability for the Compass inter-frequency carrier phase biases, and verified the tripe-frequency relative positioning with real Compass data.

PPP is an attractive positioning technique with a high accuracy using a standalone GNSS receiver, and GPS dual-frequency PPP has been an active research topic over the past decade. It is well known that static GPS-PPP is capable of providing millimeter to centimeter positioning accuracy using daily observations [7, 17]. Based on the post-processing Compass orbit and clock products from Steigenberger et al. [15], this contribution aims at showing an initial result of Compass-PPP using daily dual observations, and comparing the performance with dual-frequency GPS-PPP and combined GPS + Compass PPP. In the following sections, the un-differenced PPP algorithm for dual-frequency single and two GNSS systems is presented. Numerical results of Compass-PPP and combined GPS + Compass PPP are firstly reported and discussed.

33.2 Dual-Frequency PPP Algorithms

In the un-differenced PPP algorithm, the original GNSS code and phase observables are adopted; and LOS ionospheric delays are regarded as estimate parameters together with position, receiver clock error, zenith tropospheric delays and ambiguity. In this section, dual-frequency PPP algorithms for single and two GNSSs are both presented, including the functional and stochastic models.

33.2.1 Dual-Frequency PPP

The dual-frequency GNSS code and carrier-phase observation equations can be expressed as:

$$\begin{aligned} p^s_{r,j} &= \rho^s_r + dt_r - dt^s + T^s_r + \mu_j \cdot I^s_{r,1} + b_{r,j} - b^s_j + \varepsilon_p \\ \phi^s_{r,j} &= \rho^s_r + dt_r - dt^s + T^s_r - \mu_j \cdot I^s_{r,1} + \lambda_j \cdot M^s_{r,j} + \varepsilon_\phi \end{aligned} \tag{33.1}$$

where $p^s_{r,j}$, $\phi^s_{r,j}$ denote the code and phase observables from satellite s to receiver r on frequency j. ρ^s_r is the geometric range between satellite and receiver antennas; dt_r and dt^s refer to the receiver and satellite clock errors; T^s_r denotes the tropospheric delays; b^s_j and $b_{r,j}$ are the satellite and receiver code instrumental delays due to the transmitting and receiving hardware; $I^s_{r,1} = 40.28/f_1^2 \cdot sTEC$ is the ionospheric delay on GNSS group signal propagation at frequency f_1, $sTEC$ is the slant total integrated electron content; $\mu_j = \lambda_j^2/\lambda_1^2$, describes the dispersive

ionospheric effect, λ_j is the wavelength at frequency j; $M^s_{r,j}$ is the carrier-phase ambiguity including satellite and receiver phase instrumental delays and initial phase bias; ε_p and ε_ϕ refer to observational noise and multipath effects.

The precise satellite clock products from IGS or other organization, which always refer to ionosphere-free code and phase combinations, can be written as [5, 6]:

$$dt^s_I = dt^s + b^s_{IF} \tag{33.2}$$

where dt^s and dt^s_I denote truth and published values, the precise clock values used in this paper are from IGS and Technical University Munich (TUM, Munich, Germany) for GPS and Compass satellites respectively. $b^s_{IF} = \alpha_2 \cdot b^s_1 - \alpha_1 \cdot b^s_2$ denote the ionosphere-free combination of b^s_1 and b^s_2, $\alpha_j = \mu_j/(\mu_2 - 1)$. When the precise products are utilized, an additional combination of b^s_1 and b^s_2 is introduced to both observation equations, Eq. (33.1) can be written as:

$$\begin{aligned} p^s_{r,j} &= \rho^s_r + dt_r - dt^s_I + T^s_r + \mu_j \cdot I^s_{r,1} + b_{r,IF} + \alpha_j \cdot (B_r - B^s) + \varepsilon_p \\ \phi^s_{r,j} &= \rho^s_r + dt_r - dt^s_I + T^s_r - \mu_j \cdot I^s_{r,1} + \lambda_j \cdot M^s_{r,j} + b^s_{IF} + \varepsilon_\phi \end{aligned} \tag{33.3}$$

Similarly, $b_{r,IF} = \alpha_2 \cdot b_{r,1} - \alpha_1 \cdot b_{r,2}$ denote the ionosphere-free combination of $b_{r,1}$ and $b_{r,2}$, $B^s = b^s_2 - b^s_1$ and $B_r = b_{r,2} - b_{r,1}$ are differential code biases (DCBs) for each satellite and receiver respectively. The above-mentioned code biases can be absorbed in receiver clock, ionospheric delays and phase ambiguity, with the satellite positions X_s fixed, Eq. (33.3) can be linearized near the approximate receiver position X^0_r and the full-rank observation equations can be expressed as:

$$\begin{aligned} \Delta p^s_{r,j} &= -\,\boldsymbol{\mu}^s_r \cdot \Delta r + m^s_r \cdot \tau_r + dt^b_r + \mu_j \cdot I^{s,b}_{r,1} + \varepsilon_p \\ \Delta \phi^s_{r,j} &= -\,\boldsymbol{\mu}^s_r \cdot \Delta r + m^s_r \cdot \tau_r + dt^b_r - \mu_j \cdot I^{s,b}_{r,1} + \lambda_j \cdot M^{s,b}_{r,j} + \varepsilon_\phi \end{aligned} \tag{33.4}$$

where $\Delta p^s_{r,j}$ and $\Delta \phi^s_{r,j}$ denote the observed minus calculated observations for the code and carrier-phase, for which some systematic errors, satellite clock errors, and dry tropospheric delays, have been a priori corrected; $\boldsymbol{\mu}^s_r$ is the unit direction vector from receiver r to satellite s; and Δr denotes the three-dimensional increments of X^0_r. m^s_r and τ_r are wet tropospheric mapping function and the remaining wet zenith tropospheric delay. dt^b_r, $I^{s,b}_{r,1}$ and $M^{s,b}_r$ are the estimable receiver clock error, ionospheric delays and carrier-phase ambiguities biased by $b_{r,j}, b^s_j$, which are:

$$dt^b_r = dt_r + b_{r,IF} \tag{33.5}$$

$$I^{s,b}_{r,1} = I^s_{r,1} + \alpha_1 \cdot (B_r - B^s) \tag{33.6}$$

$$\lambda_j \cdot M^{s,b}_{r,j} = \lambda_j \cdot M^s_{r,j} - b_{r,IF} + b^s_{IF} + \alpha_j \cdot (B_r - B^s) \tag{33.7}$$

Assuming m satellites were simultaneously tracked by receiver r, then by incorporating the linearized equations (as Eq. 33.4) for all satellites, the following compact form of the observation equations can be formed:

$$\underset{4m\times 1}{y} = \underset{4m\times(3m+5)}{A} \cdot \underset{(m+5)\times 1}{x} + \underset{4m\times 1}{\varepsilon_y}\ ,\ \varepsilon_y \sim N(0, Q_{yy}) \tag{33.8}$$

$$x = [\Delta r_r, \tau_r, dt_r^b, I_{r,1}^{s,b}, M_r^{s,b}]^T \tag{33.9}$$

The unknown vector in the PPP includes three position coordinate parameters, a wet zenith tropospheric delay parameter τ_r, a receive clock correction parameter dt_r^b, slant ionospheric delays at L_1 $I_{r,1}^{s,b}$, and float ambiguity terms $M_r^{s,b}$ at each frequency, where $s = 1 \cdots m$. The quantity Q_{yy} takes the form of a diagonal matrix with its diagonal elements $Q_{ii} = \sigma_0^2 / \sin^2(E_r^s)$, E_r^s is the elevation angle of each satellite and σ_0 is the standard deviation of the GNSS observation at zenith.

33.2.2 Dual-Frequency Combined PPP

Based on Eq. (33.4), the linearized equation can be obtained after applying the GPS and Compass precise orbit and clock corrections. A systematic time difference between GPS and Compass time system is existed, that is $dt_c = dt_G + dt_{sys}$. The estimable clock corrections can be written as:

$$dt_r^{C,b} = dt_r^{G,b} + dt_{sys}^b \tag{33.10}$$

where $dt_{sys}^b = dt_{sys} + b_{r,IF}^C - b_{r,IF}^G$ is a sum of real GPS-Compass system time and a biased term $b_{r,IF}^C - b_{r,IF}^G$. Then Compass observations can be written as:

$$\begin{aligned} \Delta p_{r,j}^C &= -\boldsymbol{\mu}_r^C \cdot \Delta r + m_r^C \cdot \tau_r + dt_r^{G,b} + dt_{sys}^b + \varepsilon_P \\ \Delta \phi_{r,j}^C &= -\boldsymbol{\mu}_r^C \cdot \Delta r + m_r^C \cdot \tau_r + dt_r^{G,b} + dt_{sys}^b + \lambda_j^C \cdot M_{r,j}^{C,b} + \varepsilon_\phi \end{aligned} \tag{33.11}$$

dt_{sys}^b is the estimable GPS-Compass system time difference. The observation equations and unknown vector can be formed:

$$\underset{4m\times 1}{y} = \underset{4m\times(3m+6)}{A} \cdot \underset{(3m+6)\times 1}{x} + \underset{4m\times 1}{\varepsilon_y}\ ,\ \varepsilon_y \sim N(0, Q_{yy}) \tag{33.12}$$

$$x = [\Delta r_r, \tau_r, dt_r^b, dt_{sys}^b, I_{r,1}^{s,b}, M_r^{s,b}]^T \tag{33.13}$$

33.2.3 Dynamic Model

The Extended Kalman Filter (EKF) is utilized in the parameter estimation, and the corresponding dynamic model can be denoted as:

$$x(i+1) = \Phi_{i+1,i} \cdot x(i) + \omega(i),\ \omega \sim N(0, Q_{\omega\omega}) \tag{33.14}$$

$\Phi_{i+1,i}$ is the corresponding transition matrix, ω is the normally-distributed process noise with zero-mean and variance–covariance (VC) matrix $Q_{\omega\omega}$.

In dual-frequency PPP, the matrix $\Phi_{i+1,i}$ and $Q_{\omega\omega}$ can be denoted as:

$$\Phi_{i+1,i} = diag\{\underset{3\times3}{I}, 1, 0, \underset{m\times m}{I}, \underset{2m\times2m}{I}\} \tag{33.15}$$

$$Q_{\omega\omega} = diag\{q_p \cdot \Delta t, q_z \cdot \Delta t, q_t \cdot \Delta t, \underset{m\times m}{q_I} \cdot \Delta t, \underset{2m\times2m}{0}\} \tag{33.16}$$

where Δt is the time interval between adjacent epochs. The five sub-matrices in $\Phi_{i+1,i}$ correspond to the transition matrices of the positioning, ZTD, receiver clock error, ionospheric delays at L_1 and ambiguities; their corresponding spectral density matrices q_z, q_t and 0 indicate that these four kinds of parameters are modeled as a random walk process, white noise process, random walk process and time-invariant parameter set, respectively. The actual values of elements in the spectral density matrix depend mainly on the practical.

In the dynamic model of dual-frequency combined PPP, an additional parameter dt^b_{sys} is added in the parameter vector, with a transition matrix of identity and modeling as a time-invariant parameter.

33.3 PPP Solutions and Analysis

33.3.1 Data and Models

In this section, a comparative analysis of GPS and Compass PPP was conducted to validate the performance of Compass PPP. Three days datasets of two continuous tracking stations, named as CUT0 and CUTA, were adopted from Curtin University, Australia. The two stations were equipped with Trimble NetR9 receivers, which can track all available GNSS satellites at each open frequency, including GPS (L1, L2, and L5), GLONASS (L1, L2, and L3), Compass (B1, B2, and B3), Galileo (E1, E5a, E5b, and E5a + b) and QZSS (L1, L2, L5, and LEX) satellites. Since the two stations are near, only the results from CUT0 station are showed in this paper.

Figure 33.1 shows the sky plots of GPS and Compass constellations observed at CUT0 (32.00°S, 115.89°E) on March 20, 2012, with a cut-off elevation of 10°. For Compass system, it can be seen that four GEO satellites are stationary and always visible in the northern hemisphere, and five IGSO satellites have a track of 8-sharp. Figure 33.2 shows the visible satellites and PDOP values of GPS-only, Compass-only and combined GPS and Compass constellations at CUT0. The average number of visible satellites for Compass and combined GPS and Compass system are 7.8 and 16.5, with an increase of about 112 %. It can also be seen that there is an obvious decrease of PDOP values when Compass is added to GPS, from the average of 2.94 to 1.38. This will enhance the precision and liability of navigation service using combined GPS and Compass system.

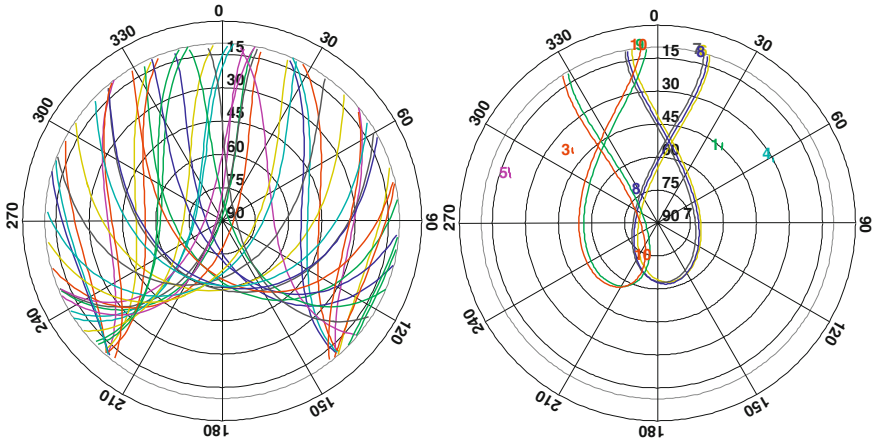


Fig. 33.1 Sky plot (azimuth vs. elevation) of GPS (*left*) and Compass (*right*) at CUT0 on March 20, 2012, cut-off = 10°

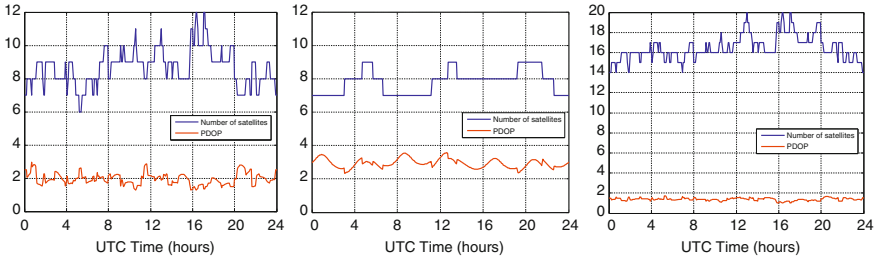


Fig. 33.2 Satellite visibility and PDOP of GPS (*left*), Compass (*middle*) and combined GPS and Compass (*right*) constellations at CUT0 on March 20, 2012, cut-off = 10°

Dual-frequency GPS and Compass observations during March 20th and 22th (DOY 80-82) 2012 are selected for PPP experiments. The Compass orbit and clock products are provided by TUM [15] in current International Terrestrial Reference Frame (ITRF) at an interval of 30 s, which is computed based on 6 Compass-capable receivers distributed in Asia–Pacific area. IGS final orbit and clock products at an interval of 15 min and 30 s, as well as the differential code bias products are used for GPS satellite corrections. For all PPP solution, a cut-off elevation of 10° is adopted and UNB3m model and Global Mapping Function (GMF) are utilized for priori correction of tropospheric delay and mapping function of ZTD estimation. The filter convergence time is defined when the positioning accuracy of each component is better than 10 cm. The reference coordinates for CUT0 and CUTA station are obtained by network adjustment using Bernese 5.0 software, with an accuracy of 3 mm for horizon and 5 mm for vertical.

33.3.2 Dual-Frequency PPP

Compass B1 and B2 frequency observations for three days are utilized for static Compass-PPP. The positioning results are shown in Fig. 33.3, together with the statistics of results after convergence in Table 33.1.

From Fig. 33.3 and Table 33.1, it can be concluded that, based on the existing Compass products from TUM, the positioning results of 1–3 cm can be achieved in dual-frequency Compass-PPP. After filter convergence, the 3D positioning accuracy is better than 3 cm; and the stability of filter can be achieved at a level of 1–2 cm.

The corresponding dual-frequency GPS-PPP results and the statistics are showed in Fig. 33.4 and Table 33.2 respectively. As we know, PPP solutions are sensitive to satellite orbit and clock products. Currently, the Compass orbit and clock products from TUM have a lower accuracy than GPS products from IGS. Therefore, the convergence time of Compass-PPP is longer than GPS-PPP. After filter convergence, the positioning accuracy of Compass-PPP is comparative to GPS-PPP, both at a level of cm.

Figure 33.5 shows the difference of zenith tropospheric delays (ZTD) and receiver clock error estimations from GPS-only and Compass-only PPP. The difference of ZTD is less than 2 cm, with a bias of 0.8 cm and standard deviation of 1.0 cm. Based on Eq. (33.5), the difference of receiver clock estimations can be expressed as:

$$\left(dt_r^b\right)_C - \left(dt_r^b\right)_G = dt_{sys} + b_{r,IF}^C - b_{r,IF}^G \tag{33.17}$$

From Eq. (33.17), it is obvious that above term is same as the estimable biased GPS-Compass time difference dt_{sys}^b, which can be regarded as a constant in daily data; Fig. 33.5 shows that the difference of receiver clock error tends to 571.3–571.5 m, with a standard deviation of 0.1 m.

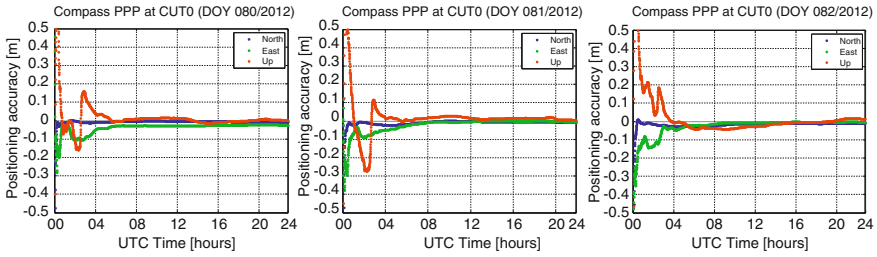


Fig. 33.3 Dual-frequency Compass-PPP positioning results for three days at CUT0

Table 33.1 Statistics of Compass-PPP solutions after convergence at CUT0, in cm

Component	DOY80			DOY81			DOY82		
	Bias	STD	RMS	Bias	STD	RMS	Bias	STD	RMS
North	−0.5	0.2	0.5	−0.6	0.6	0.9	−1.5	0.6	1.6
East	−2.7	0.6	2.8	−0.8	1.3	1.5	−1.0	1.0	1.4
Up	0.5	1.1	1.2	1.5	0.7	1.7	−1.5	1.8	2.4

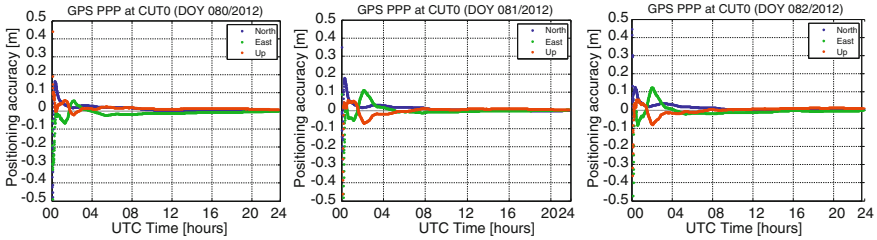


Fig. 33.4 Dual-frequency GPS-PPP positioning results for three days at CUT0

Table 33.2 Statistics of GPS-PPP solutions after convergence at CUT0, in cm

Component	DOY80			DOY81			DOY82		
	Bias	STD	RMS	Bias	STD	RMS	Bias	STD	RMS
North	0.7	0.8	1.1	0.6	0.8	1.0	0.6	0.7	0.9
East	−1.0	0.6	1.2	−0.2	0.5	0.6	−0.9	0.6	1.1
Up	1.3	0.3	1.4	0.5	0.7	0.9	0.5	0.8	0.9

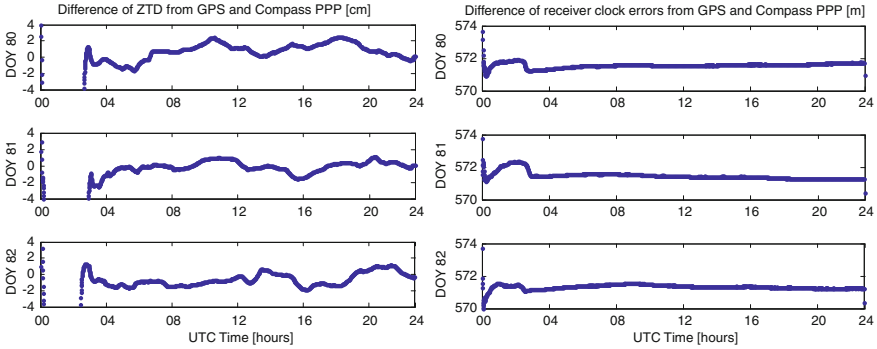


Fig. 33.5 Difference of ZTD (*left*) and receiver clock error (*right*) estimations from GPS and Compass PPP

33.3.3 Dual-Frequency Combined PPP

To evaluate the PPP performance using single and combined GNSS systems, GPS and Compass observations are integrated to perform combined PPP. The initial STD values for GPS and Compass code and phase observations are vital important to define the measure precision of code and phase, as well as their relative precision, a proper weight ratio between GPS and Compass observations can shorten the convergence time and improve the filter stability. Figure 33.6 shows positioning mean and STD of the combined PPP with different STD ratios, where ratio $= \mathrm{STD_C}/\mathrm{STD_G}$.

Figure 33.6 shows that when we give an equal weight for GPS and Compass observations, the stability and accuracy of positioning results will achieve 2–4 and 3.5–4.5 cm respectively, and the optimal STD ratio for Compass/GPS observations is 2.2. Figure 33.7 shows the combined PPP results on DOY080 and 081/2012 with a weight ratio of 2.2. The results show that, combined GPS + Compass PPP can also achieve a positioning accuracy of cm level, and it can shorten the convergence time, from 33 to 30 min, but not necessarily improve positioning results by much if the satellites of the single GNSS system already have good receiver-satellite geometry.

The local overall model (LOM) test analyses the discrepancy between data and given model by means of least squares residual vector $\hat{e}$, with an average of 1 [12]. Where $LOM = \hat{e}^T Q^{-1} \hat{e} / red$, when applying the LOM test in EKF, $\hat{e}$ is the predicted residual, Q is the predicted covariance matrix, and red represents the redundancy at the current epoch. Figure 33.8 shows the LOM test statistics accompanying PPP results for Compass and GPS respectively. The repeatability of LOM test for Compass PPP due to the multipath effect.

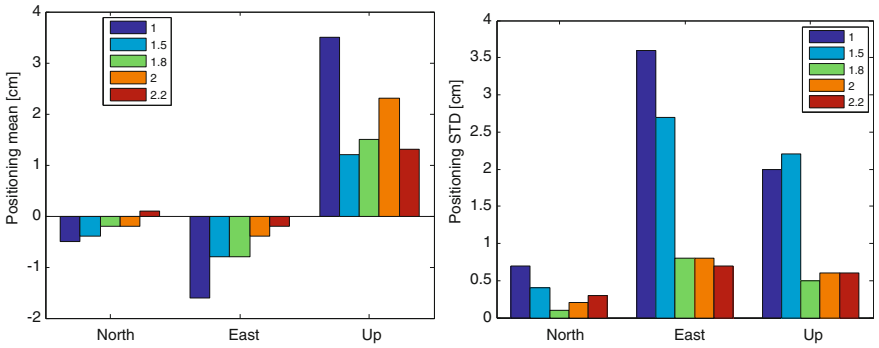


Fig. 33.6 Positioning mean and STD of combined PPP with different STD ratios for Compass/GPS observations

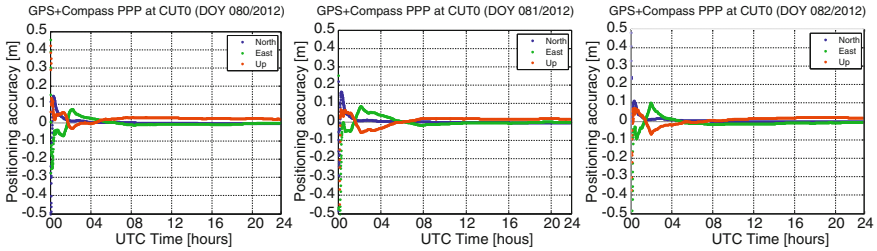


Fig. 33.7 Combined GPS + Compass PPP positioning results for two days at CUT0, STD ratio = 2.2

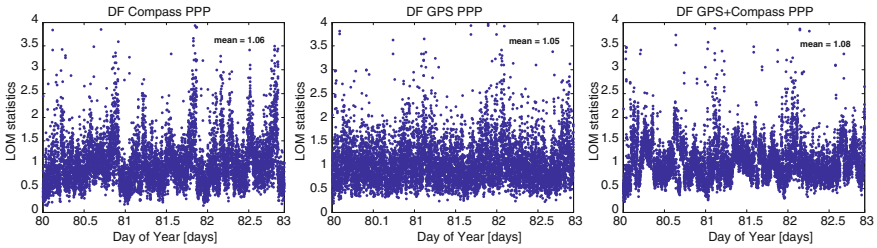


Fig. 33.8 LOM values for Compass, GPS and combined PPP, DOY80-82/2012

33.4 Conclusions

This study reports the performance of dual-frequency Compass-PPP and combined GPS + Compass in static mode in comparison with GPS-PPP. Based on the Compass products from TUM, three daily datasets have been used in the numerical analysis. The main conclusions are as followed:

1) Currently, dual-frequency Compass-PPP is capable of providing static positioning at centimeter level, which is comparative to GPS-PPP.
2) Combined GPS + Compass PPP shorten the convergence time slightly, but it contributes only marginally to the GNSS-only positioning accuracy.

It should be noted that the Compass orbit and clock products from TUM are only based on six stations, and have a lower accuracy than IGS products, which result in the longer time for convergence and less positioning accuracy of Compass-PPP. In future, with more accuracy of Compass products, the positioning accuracy of Compass-PPP will be further improved. Furthermore, PPP technique using Compass triple-frequency signals also needs for attention.

Acknowledgments Thanks to Technology University of Munich (TUM) and German Aerospace Center (DLR) for providing the Compass orbit and clock products. Thanks to Dr. Andrea Nardo from Curtin University for providing the reference coordinates of CUT0 and CUTA stations. This work was jointly supported by National Key Program 863(Grant No. 2011AA120503), National

Key Technology R&D Program (Grant No. 2012BAB16B01), the Fundamental Research Funds for the Central Universities (Grant No. 201121402020009) and Key Laboratory of Geospace Environment & Geodesy Ministry of Education project (Grant No.11-01-08) for financial support.

References

1. Cao C, Jing G, Luo M (2008) COMPASS satellite navigation system development. PNT challenges and opportunities symposium, Stanford, California, USA, 5–6 Nov 2008
2. Cao W, O'Keefe K, Cannon ME (2008) Evaluation of COMPASS ambiguity resolution performance using geometric-based techniques with comparison to GPS and Galileo. ION GNSS 2008, Session B4, Savannah, Georgia, USA, 16–19 Sept 2008
3. Chen HC, Huang YS, Chiang KW, Yang M, Rau RJ (2009) The performance comparison between GPS and BeiDou-2/COMPASS: a perspective from Asia. J Chin Inst Eng 32(5):679–689
4. Chen J, Wu B, Hu X, Zhou S, Cao Y, Wu X, Xing N (2012) Compass/Beidou: system status and initial service; IGS workshop on GNSS Biases, University of Bern, Switzerland, 18–19 Jan 2012
5. Dach R, Brockman E, Schaer S et al (2009) GNSS processing at CODE: status report. J Geodesy 83:353–365. doi:10.1007/s00190-008-0281-2
6. Dow J, Neilan R, Rizos C (2009) The international GNSS service in a changing landscape of global navigation satellite systems. J Geodesy 83:191–198. doi:10.1007/s00190-008-0300-3
7. Kouba J, Héroux H (2001) Precise point positioning using IGS orbit and clock products. GPS Solutions 5(2):12–28
8. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Riley S (2012) A COMPASS for Asia: first experience with the BeiDou-2 regional navigation system. IGS Workshop 2012, Olsztyn, 23–27 July 2012
9. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen PJG, Nakamura S (2012) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. GPS Solutions. doi:10.1007/s10291-012-0272-x
10. Nadarajah N, Teunissen PJG, Buist P and Steigenberger P (2012) First results of instantaneous GPS/Galileo/COMPASS attitude determination. In: Proceedings of the 6th ESA workshop on satellite navigation user equipment technologies (NAVITEC'12)
11. Nadarajah N, Teunissen PJG and Raziq N (2012) Instantaneous COMPASS-GPS attitude determination: a robustness analysis. Advances in Space Research (submitted)
12. Teunissen PJG (1990) Quality control in integrated navigation systems. IEEE Aerosp Electron Syst Mag 5(7):35–41
13. Shi C, Zhao QL, Li M, Tang WM, Hu ZG, Lou YD, Zhang HP, Niu XJ, Liu JN (2012) Precise orbit determination of Beidou satellites with precise positioning. Sci China Earth Sci 55(7):1079–1086
14. Shi C, Zhao QL, Hu ZG, Liu JN (2012) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. GPS Solutions. doi:10.1007/s10291-012-0264-x
15. Steigenberger P, Montenbruck O, Hauschild A, Hugentobler U (2012) Performance analysis of compass orbit and clock determination and compass-only PPP. IGS Workshop 2012, Olsztyn, 23–27 July 2012
16. Yang YX, Li JL, Xu JY, Tang J, Guo HR, He HB (2011) Contribution of the Compass satellite navigation system to global PNT users. Chin Sci Bull 56(26):2813–2819. doi:10.1007/s11434-011-4627-4
17. Zumberge J, Heflin M, Jefferson D, Watkins M, Webb F (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102(B3):5005–5017

Chapter 34
Fixed Failure Rate Ambiguity Validation Methods for GPS and Compass

Lei Wang and Yanming Feng

Abstract Ambiguity resolution plays a crucial role in real time kinematic GNSS positioning which gives centimetre precision positioning results if all the ambiguities in each epoch are correctly fixed to integers. However, the incorrectly fixed ambiguities can result in large positioning offset up to several meters without notice. Hence, ambiguity validation is essential to control the ambiguity resolution quality. Currently, the most popular ambiguity validation is ratio test. The criterion of ratio test is often empirically determined. Empirically determined criterion can be dangerous, because a fixed criterion cannot fit all scenarios and does not directly control the ambiguity resolution risk. In practice, depending on the underlying model strength, the ratio test criterion can be too conservative for some model and becomes too risky for others. A more rational test method is to determine the criterion according to the underlying model and user requirement. Miss-detected incorrect integers will lead to a hazardous result, which should be strictly controlled. In ambiguity resolution miss-detected rate is often known as failure rate. In this paper, a fixed failure rate ratio test method is presented and applied in analysis of GPS and Compass positioning scenarios. A fixed failure rate approach is derived from the integer aperture estimation theory, which is theoretically rigorous. The criteria table for ratio test is computed based on extensive data simulations in the approach. The real-time users can determine the ratio test criterion by looking up the criteria table. This method has been applied in medium distance GPS ambiguity resolution but multi-constellation and high dimensional scenarios haven't been discussed so far. In this paper, a general ambiguity validation model is derived based on hypothesis test theory, and fixed failure rate approach is introduced, especially the relationship between ratio test threshold and failure rate is examined. In the last, factors that influence fixed failure rate

L. Wang (✉) · Y. Feng
Science and Engineer Faculty, Queensland University of Technology, Queensland, Australia
e-mail: l62.wang@qut.edu.au

L. Wang
School of Geodesy and Geomatics, Wuhan University, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_34,

approach ratio test threshold is discussed according to extensive data simulation. The result shows that fixed failure rate approach is a more reasonable ambiguity validation method with proper stochastic model.

Keywords Ambiguity validation · Ratio test · Fixed failure rate · Quality control · Multi-constellation

34.1 Introduction

Integer ambiguity resolution is the key to fast and high precision GNSS positioning and navigation. With correctly fixed ambiguity, centimetre to millimetre positioning accuracy is achievable, incorrectly fixed ambiguity may result unacceptable positioning errors. Hence, ambiguity validation is very crucial to control positioning result quality.

Ambiguity validation problem has been an important issue in GNSS positioning research and applications. For a long time, the problem was mainly addressed by discrimination test, Frei constructed F-ratio test [2, 6], Euler simplified the form and recommeded to use 'Ratio test' instead [1], Tiberius used difference test and gave an empirical threshold for difference test [25], Han and Wang used student's test instead of F-test and constructed a w-statistic for ambiguity validation propose [3, 32]. Teunissen proposed integer aperture estimation and gave a unified framework for ambiguity validation and ambiguity estimation [18]. Since then, five new integer aperture estimators was proposed, including ellipsoidal integer aperture [17], integer aperture bootstrapping [21] ,integer aperture least-squares [22], penalized integer aperture [19], optimal integer aperture estimator [20]. Later on, fixed-failure ratio test are proposed as a improvement of fixed threshold ratio test [23, 28], A comparison between different ambiguity validation methods was made by Verhagen and Li [7, 27], Ji tried to combine ellipsoidal integer aperture estimator and ratio test for ambiguity validation [5]. A variable duration fixed failure rate ambiguity validation method was examined [11]. A further study on fixed-failure rate ratio test study is carried out by Verhagen [31].

Popular discrimination tests focused on statistic construction, these statistics all based on assumption of normal distribution of real-valued ambiguity, but what they real test is ambiguity residual. Unfortunately, ambiguity residual is not normally distributed. Another important but never been studied topic about fixed failure rate is the implicit relationship between hypothesis test threshold and failure rate. Moreover, to the best of author's knwoledge, the factors that impact the threshold of hypothesis test haven't been systematically studied. Hence, the contribution of this paper can be summarised as:

1. Investigation of the ambiguity validation with the hypothesis test theory, and the relationship between general hypothesis test and ambiguity validation

problem. The general ambiguity validation model based on the distribution of ambiguity residual is derived.
2. Examination of the implicit relationship between failure rate and validation threshold. Usually, given validation threshold, the failure rate can be statistically calculated. Finding the relationship between failure rate and threshold is a reverse problem, and the problem is solved by curve fitting method which gives an intuitive impression about the relationship. The relationship is the key to fixed failure rate approach, but never been systematically analysed.
3. Understanding the factors impacting validation threshold through extensive numerical simulation. Simulation results can help to explore how ambiguity dimension, frequency number, and prior stochastic model would affect the threshold of hypothesis test in the fixed failure rate approach.
4. Analysis of the fixed failure rate ambiguity validation method in a multi-constellation environment. In particular,investigation focuses on how the fixed failure rate test approach would perform in the high dimension case for the current GPS and Compass constellations. The experiment results should especially benefit GPS and Compass users.

The following sections are organized as follows: Ambiguity resolution model and ambiguity residual distribution are briefly reviewed in Sect. 34.2. In the Sect. 34.3, general hypothesis test theory is mentioned first. Ambiguity validation problem is treated as a special hypothesis test case. The most popular discrimination test, ratio test is taken as an example to demonstrate the existence of relationship between failure rate and hypothesis test threshold. In Sect. 34.4, A comprehensive simulation is carried out to investigate the relationship between frequency number, ambiguity dimension and prior stochastic model. Finally, it comes the conclusion section.

34.2 Probability Distribution of Ambiguity Residual

All hypothesis tests are based on a distribution and a test statistic, before we move to further discussion, the distribution of population should be investigated. For ambiguity validation problem, the population is ambiguity residual, the defination and distribution of ambiguity residual will be discussed in this section.

34.2.1 Definition of Ambiguity Residual

The function model for GNSS positioning can be expressed as a mixed integer estimation problem:

$$y = Aa + Bb + e, Q_y = \sigma_0^2 P^{-1} \tag{34.1}$$

where y is the given GNSS observation, a and b are integer parameters and real-valued parameters respectively. A and B are corresponding design matrices. e denotes noise vector. σ_0 is prior unit weight standard deviation. P standards for weight matrix and Q_y is the variance matrix of observations y, generally, observations are considered being independent on each other.

A typical ambiguity resolution procedure follows three steps: the first step is applying standard least-squares, the 'float' solution can be obtained and denoted as $\hat{a}$ and $\hat{b}$. In this stage, the integer nature of ambiguity parameters is simply discarded. Assuming that the observation vector y follows normal distribution, the estimates $\hat{a}$ and $\hat{b}$ follow normal distribution as well, the corresponding variance-covariance matrix reads:

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \begin{pmatrix} Q_{\hat{a}} & Q_{\hat{a}\hat{b}} \\ Q_{\hat{b}\hat{a}} & Q_{\hat{b}} \end{pmatrix} \tag{34.2}$$

The second phase is mapping the real-valued ambiguity parameter $\hat{a}$ to integers with integer estimator,which is also known as integer ambiguity estimation. The optimal integer estimator is integer least-squares [14], and the objective function of integer least-squares is given as:

$$\min(\hat{a} - z)Q_{\hat{a}}^{-1}(\hat{a} - z), \quad z \in Z^n \tag{34.3}$$

Once $\hat{a}$ been fixed to integer, Corresponding ambiguity residuals can be defined as [12, 30]:

$$\check{\epsilon} = \hat{a} - \check{a}, \hat{a} \in R^n, \quad \check{a} \in Z^n \tag{34.4}$$

where $\check{a}$ is fixed integer ambiguity.

The third step is to update real-valued parameters $\hat{b}$, which follows:

$$\check{a} = \hat{a} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}^{-1}(\hat{a} - \check{a}) \tag{34.5}$$

An integer estimator is the key to map the real-value points to their integer points. The set of real-valued points closer to a certain integer z than any other integer points are called 'pull-in region' S_z [15]. (or Voronoi cells [4, 13]), the definition of pull-in region reads:

$$S_z = x \in R^n | z = S(x), \quad z \in Z^n \tag{34.6}$$

where the subset S_z denotes a pull-in region centred at z. All $\hat{a}$ falls in S_z will be mapped to the $z \in Z^n$ [13]. The integer ambiguity estimation procedure can be described as:

$$\check{a} = \sum_{z \in Z^n} z * s_z(\hat{a}), s_z(x) = \begin{cases} 1 & \text{if } x \in S_z \\ 0 & \text{otherwise} \end{cases} \tag{34.7}$$

where s_z denotes the indicator function of the pull-in region S_z. The shape of pull-in region S_z depends on integer estimator. The integer estimator satisfy following three conditions are called admissible integer estimator [14]:

$$\begin{cases} \bigcup\limits_{z\in Z^n} S_z = R^n \\ S_{z_1} \bigcap S_{z_2} = \emptyset, \forall z_1 \neq z_2, z_1 \in Z^n, \quad z_2 \in Z^n \\ S_z = z + S_0, \quad z \in Z^n \end{cases} \tag{34.8}$$

Particularly, the third condition in Eq. 34.8 shows S_z is insensitive to integer z, which is known as 'z-translational invariant'. According to the definition of the ambiguity residual, the difference between $\hat{a}$ and $\check{\epsilon}$ is $\breve{a}$, while $\breve{a}$ is an integer vector. $\check{\epsilon}$ can be viewed as relative position of $\hat{a}$ respect to $\breve{a}$, and $\hat{a}$ can be viewed as the absolute position of $\hat{a}$ in R^n. Obviously, ambiguity residual $\check{\epsilon}$ is more important than $\breve{a}$ in ambiguity estimation and validation.

34.2.2 Distribution of Ambiguity Residual

The GNSS observations y are assumed as normal distributed. According to the variance propagation law, the 'float solution' $\hat{a}$ follows normal distribution as well, the probability density function (PDF) of $\hat{a}$ can be written as:

$$f_{\hat{a}}(x) = \frac{1}{\sqrt{detQ_{\hat{a}}}(2\pi)^{\frac{1}{2}n}} \, exp\left\{-\frac{1}{2}\|x-a\|^2_{Q_{\hat{a}}}\right\} \tag{34.9}$$

where a is the correct integer vector. Similar to $\hat{a}$, distribution of $\check{\epsilon}$ can be described by PDF as well. According to definition of ambiguity residual, the PDF of $\check{\epsilon}$ can be given as [16, 30]:

$$f_{\check{\epsilon}}(x) = \sum_{z\in Z^n} f_{\hat{a}}(x+z)s_0(x), x \in R^n \tag{34.10}$$

Applying 'z-translational invariant' property, the pull-in region can be transformed from S_a to S_0. $f_{\check{\epsilon}}(x)$ equals to infinite sum of normal distribution PDF, and equation $\int_{s_0} f_{\check{\epsilon}}(x)dx = 1$ always holds true. Equation (34.10) can be decomposed into two parts as:

$$f_{\check{\epsilon}}(x) = f_{\hat{a}}(x)s_0(x) + \sum_{z\in Z^n\setminus\{0\}} f_{\hat{a}}(x+z)s_0(x), x \in R^n \tag{34.11}$$

Within the pull-in region, the first term on the right of Eq. (34.11) is exactly the same as PDF of $\hat{a}$. For 1-dimensional case, the relationship between $\check{\epsilon}$ and $\hat{a}$ are shown in Fig. 34.1. According to Eq. (34.11), $f_{\check{\epsilon}}(x) \geq f_{\hat{a}}(x)$ always holds true.

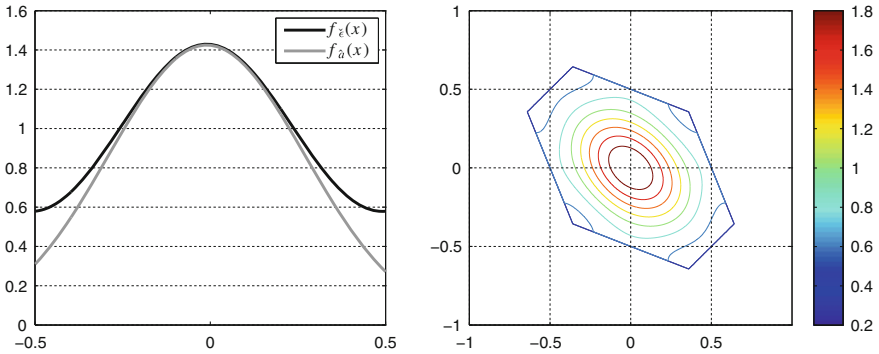


Fig. 34.1 Demonstration of $f_{\check{\epsilon}}(x)$ and $f_{\hat{a}}(x)$ for 1-D case (*left*) and 2-D case (*right*). $\hat{a}$ follows normal distribution, $f_{\check{\epsilon}}(x)$ depends on not only $Q_{\hat{a}}$, but also integer estimator, the *right* sub figure demonstrates $f_{\check{\epsilon}}(x)$ based on integer least-squares

It is noticed that $\check{\epsilon}$ is defined in S_0 and S_0 is defined by an integer estimator. For different integer estimator, S_0 is different. Hence, $f_{\check{\epsilon}}(x)$ not only depends on $f_{\hat{a}}(x)$, but also integer estimators. For 1-dimensional case, integer rounding, integer bootstrapping and integer least-squares are identical, hence, $f_{\check{\epsilon}}(x)$ is same for all three integer estimators. However, this conclusion doesn't holds true for any correlated higher dimensional case, because S_0 for different integer estimators are different in correlated high dimensional case. $\hat{a}$ follows multivariate normal distribution, $f_{\hat{a}}(x)$ can be described by expectation and $Q_{\hat{a}}$. $\check{\epsilon}$ follows irregular distribution, but $f_{\check{\epsilon}}(x)$ is linear combination of $f_{\hat{a}}(x)$. Once S_0 is given, $f_{\check{\epsilon}}(x)$ can be uniquely defined by $f_{\hat{a}}(x)$.

34.3 Hypothesis Test for Ambiguity Validation Problem

Ambiguity validation problem is a hypothesis test problem. All existing ambiguity validation methods are hypothesis tests, regardless of discrimination test or integer aperture estimation. From this perspective, all ambiguity validation methods can be unified into the integer aperture framework. Particularly, ambiguity validation problem is distinguished because of two special features: (1) regarding the population of hypothesis test, ambiguity residual is not normal distributed; (2) the sample space is confined by pull-in region S_0. The existing ambiguity validation research efforts have focused on statistic construction. In this context, a ambiguity validation problem is treated as general hypothesis test problem.

Table 34.1 Hypothesis test decision matrix

		Reality	
		H_0	H_a
Statistical decision	'H_0'	$1-\alpha$	β
	'H_a'	α	$1-\beta$

34.3.1 Review of General Hypothesis Test Theory

For any n-dimensional stochastic variables x, each sample can be presented as a n-dimensional coordinate, denoted as a sample point. All possible sample points lay in the sample space, denoted by E. The probability of occurrence of every sample point is presented by probability density function (PDF). If the PDF varies continuously over the sample space, a continuous region within the sample space that satisfy a particular hypothesis H_0 can always be identified [9], the region is known as acceptance region Ω:

$$\Omega(x|H_0) \subset E, x \in R^n \tag{34.12}$$

Similarly, the reject region can be defined. The probability of sample points that fall in the acceptance region can be calculated by the integral of PDF over the acceptance region. The integral is known as success rate. Theoretically, the PDF continuously varies over R^n for a n-dimensional normally distributed stochastic variables. In this case, there is no rigorous boundary between null hypothesis H_0 and alternative hypothesis H_a. However, the output of hypothesis tests can only be either accept or reject. Hence, it is almost impossible to find an acceptance region for H_0 with zero risk. The relationship between reality and hypothesis test decision can be shown in Table 34.1 [10].

According to Table 34.1, H_0 and H_a means true and false respectively in reality, while H_0 and H_a means accept and reject respectively in decision. Probability of false reject H_0 is denoted by α, which is also known as probability of 'Type I error'. Probability of false accept H_a is denoted by β, which is also known as 'type

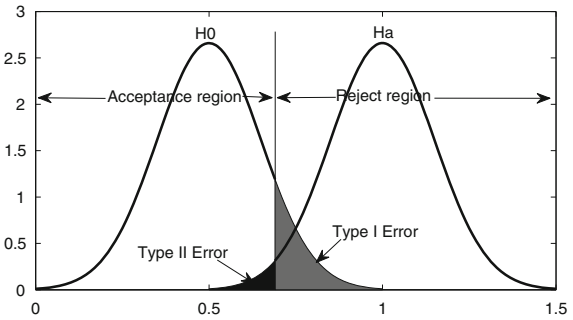


Fig. 34.2 Hypothesis test acceptance region, reject region, type I error and type II error

II error'. For many hypothesis problems, the population follows normal distribution, The risk of type I error and type II error are illustrated in Fig. 34.2.

34.3.2 Hypothesis Test Model for Ambiguity Validation

The existing ambiguity validation methods base on the normal distribution of $\hat{a}$, but actual statistics base on ambiguity residual. Obviously, it is not reasonable. Hypothesis test should base on ambiguity residual distribution. In this section, a hypothesis test model for ambiguity validation problem is derived from ambiguity residual distribution and hypothesis test theory. The purpose of hypothesis test is just making a good decision subject to given distribution.

According to the definition of ambiguity residual PDF, $f_{\check{\epsilon}}(x)$ is an infinite sum of $f_{\hat{a}}(x)$, $f_{\hat{a}}(x)$ follows a normal distribution. Moreover, $f_{\hat{a}}(x)$ is continuous over the R^n for an n-dimensional ambiguity validation problem. The ambiguity residual distribution $f_{\check{\epsilon}}(x)$ is linear combination of $f_{\hat{a}}$, hence, $f_{\check{\epsilon}}(x)$ is continuous over the sample space as well. Each ambiguity validation methods corresponds to a particular acceptance region, while the shape of acceptance region is determined by the hypothesis test and the size of acceptance region is determined by the hypothesis test threshold.

The hypothesis test for ambiguity validation can be described as:

$$H_0 : \check{a} = 0, H_a : \check{a} \neq 0 \tag{34.13}$$

where $\hat{a}$ is the output of integer estimator. As the distribution of ambiguity residual is 'z-translational invariant', assumming true integer ambiguity vector is 0. Regarding the distribution of ambiguity residual (seeing Eq. (34.11)). In Eq. (34.11), the first term on the right of corresponds to the probability of H_0 and the second term corresponds the probability of H_a. The hypothesis test model for ambiguity validation can be expressed in Fig. 34.3. Sample space E for ambiguity validation is S_0, particularly, S_0 specifies the interval $[-0.5, 0.5]$ for 1-dimension case. Integer estimator is a special hypothesis test as well, the 'pull-in region' of integer estimator is corresponding acceptance region. For integer estimators, every

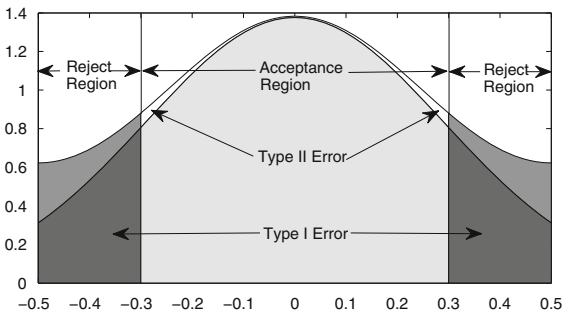


Fig. 34.3 Illustration of 1-Dimensional hypothesis test model for ambiguity validation problem. *Upper curve* indicate $f_{\check{\epsilon}}(x)$ and *lower curve* indicate $f_{\hat{a}}(x)$. *white* region between two curve lines stands for type II error while the *dark grey* region stands for type I error

sample falls in the sample space will be accepted. Hence, additional test is required to exclude a portion of samples for quality control propose. integer estimator is a special case of ambiguity validation method, which is consistency with integer aperture framework [18].

According to Fig. 34.3, it is obvious that the hypothesis test criterion determine acceptance region as well as the probabilities of type I error and type II error. The probability of type I error and type II error are known as false alarm rate and failure rate respectively. Probability of correctly acceptance is called success rate. Similar to success rate of integer estimator, success rate for ambiguity validation can be calculated by:

$$P_s = \int_{\Omega} f_{\hat{a}}(x)dx, \Omega \in S_0 \tag{34.14}$$

Corresponding failure rate for ambiguity validation can be given as:

$$P_f = \int_{\Omega} (f_{\check{\epsilon}}(x) - f_{\hat{a}}(x))dx, \Omega \in S_0 \tag{34.15}$$

where Ω is acceptance region, x is the position of $\check{\epsilon}$. With given $f_{\check{\epsilon}}(x)$ and hypothesis test, a larger acceptance region results in larger failure rate and vice verse. Hence, the failure rate can be controlled by chosen proper hypothesis test threshold for any hypothesis test. A negative effect of low failure rate is that it results in large false alarm rate.

The characteristic of ambiguity residual distribution should be systematically analysed before moving forward. In order to reveal the characteristic of $f_{\check{\epsilon}}(x)$, a indicator can be constructed as:

$$\delta(x) = \frac{f_{\hat{a}}(x)}{f_{\check{\epsilon}}(x)} \tag{34.16}$$

$\delta(x)$ indicates the proportion of $f_{\hat{a}}(x)$ subject to $f_{\check{\epsilon}}(x)$ at x. 1 dimensional $\delta(x)$ for different $Q_{\hat{a}}$ is shown in Fig. 34.4 [24]. Two conclusions can be drawn from Fig. 34.4: (1) $\delta(x)$ is different for different $f_{\hat{a}}(x)$ or $f_{\check{\epsilon}}(x)$. For the same x, proportion of $f_{\hat{a}}(x)$ for stronger model is always larger than for weak model, or x for stronger model always have lower failure risk respect to weaker model. The model strength indicated by variance, larger variance means weaker model; (2) even for strong model, x close to boundary of pull-in region still unreliable, it is the necessity of ambiguity validation. For example, if $\sigma = 0.1$ and confidence interval is 3σ, that means about 99.7 % of $\hat{a}$ will fall in $[-0.3, 0.3]$ theoretically. If the $\hat{a}$ close to 0.5 or -0.5, corresponding $f_{\hat{a}}\hat{a}$ subject to $f_{\check{\epsilon}}(\hat{a})$ is smaller than 50 %, which means although the $\hat{a}$ fall in S_0 and integer estimator can fix $\hat{a}$ to 0, but the probability of correctly fixed is less than 50. For multi-dimensional case, a similar conclusion can be obtained, as $f_{\hat{a}}$ following multi-variate normal distribution, which involves only one peak in the pull-in region, and $f_{\hat{a}}(x)$ is continuous over the pull-in region. $f_{\check{\epsilon}}(x)$ is infinite sum of $f_{\hat{a}}(x)$, which have a similar property with $f_{\hat{a}}$. Hence, one

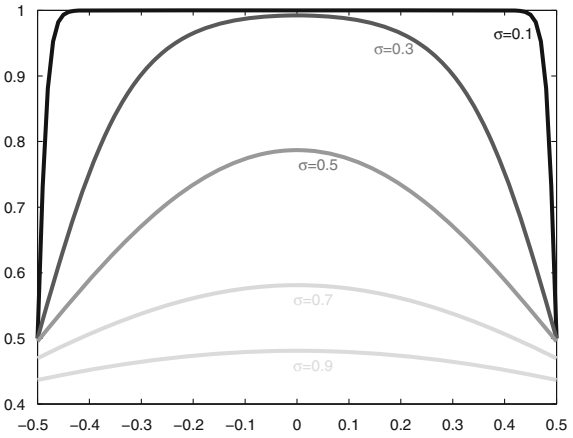


Fig. 34.4 1-Dimensional $\frac{f_{\hat{a}}(x)}{f_{\check{a}}(x)}$ for different variance

acceptance region can be found for every hypothesis test in a multivariate case. One to one map between failure rate and hypothesis test criterion is thus achieved.

34.3.3 Fixed Failure Rate Approach and Fixed Failure Rate Ratio Test

Having examined ambiguity residual distribution, the remaining work is to construct a hypothesis test. There are many existing ambiguity validation methods, such as F-ratio test, ratio test, difference test, projector test, ellipsoidal integer aperture, integer aperture bootstrapping, integer aperture least-squares, penalized integer aperture, optimal integer aperture et al. The propose of ambiguity validation is to ensure the reliability of fixed ambiguities. Hence, type II error, which is acceptance incorrect integer ambiguity, is of great concern. According to Eq. (34.15), failure rate is a function of acceptance region. Acceptance region includs two aspects: shape and size. Shape is determined by the statistic of hypothesis test, Verhagen examined acceptance region for all test enumerated above [27, 29]. Acceptance region size is determined by hypothesis test threshold. Hence,once hypothesis is chosen, failure rate is a function of hypothesis test threshold. As mentioned, the negative effects of smaller failure rate is larger type I error probability. Although type I error is not as critical as type II error in terms of risk, it is still necessary to keep it not too large. The decision should be made according to the user's demand and achieving a good balance between type I error and type II error. That is the reason why the fixed failure rate approach is recommended rather than a 'failure free' approach. The fixed failure rate approach is applicable to any hypothesis test, but different hypothesis test may lead to different performance.

As discussed, a failure rate can be calculated by integration of type II error probability over the acceptance region. Usually, failure rate is calculated after the

hypothesis test threshold is given. Finding hypothesis test threshold from a given failure rate is a reverse problem. The specific relationship between the hypothesis tet threshold and failure rate depends on specific hypothesis test. In order to explain the relationship clearly,the most popular ambiguity validation method, ratio test is used as an example. The definition of ratio test can be written as [1, 27] :

$$\frac{\|\hat{a}-\check{a}\|^2_{Q_{\hat{a}}}}{\|\hat{a}-\check{a}'\|^2_{Q_{\hat{a}}}} \leq \mu \tag{34.17}$$

If $\check{a}$ satisfy the Eq. (34.17), $\check{a}$ is accepted as correct integer vector. Here the reciprocal form of ratio test is used,because in the reciprocal form, μ varies between 0 and 1. For each μ, acceptance region can be identified. In the 2-dimensional case ratio test acceptance regions for different μ are demonstrated in Fig. 34.5. As shown in this figure, a larger μ result in a larger acceptance region. For the construction procedure of ratio test pull-in region, we refer to Verhagen [27, 29], Similar acceptance region can be identified for all ambiguity validation methods and different hypothesis tests have different acceptance region shape.

Referring to ratio test example, procedure of the fixed failure rate ratio test (FF-ratio test) involves four steps:

1. Simulation. For a given float solution $\hat{a}$ and $Q_{\hat{a}}$, a number of 'float solution' samples $\tilde{a}$ following N (0,$Q_{\hat{a}}$) distribution can be generated, N(·) standards for normal distribution.
2. Calculation. Trying to fix every simulated samples $\tilde{a}$ with LAMBDA method, and statistic incorrectly fixed sample number and corresponding ratio test value.
3. Finding proper threshold. Set up a relationship between failure rate and ratio test value, finding ratio test threshold μ for the specific failure rate.
4. Validate real data with fixed failure rate ratio test threshold. Once fixed failure rate ratio test threshold μ is acquired, the ratio test value for real data $\hat{a}$ should

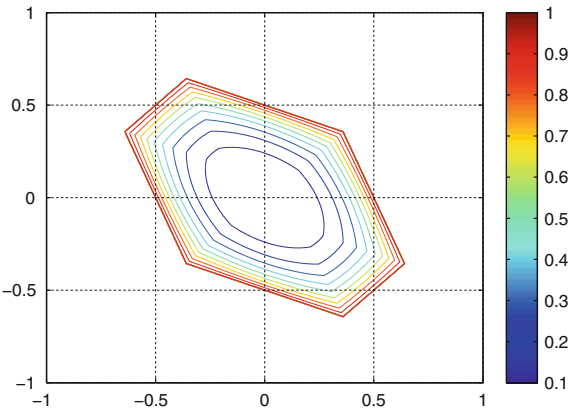


Fig. 34.5 A demonstration of ratio test acceptance region for different μ (2-D case)

be calculated. Then compare μ and ratio test value of $\hat{a}$ to decide accept the fixed integer ambiguity or not.

The first step defines the normal distribution of $\hat{a}$, the second step defines the ambiguity residual distribution $\check{\epsilon}$, Because integer least-square is adopted as integer estimator, the corresponding $f_{\check{\epsilon}}(x)$ respect to $Q_{\hat{a}}$ and integer least-squares pull-in region. It is not necessary to use integer least-squares as integer estimator, integer bootstrapping and integer rounding are admissible integer estimators as well, but integer least-square is more popular because it is the optimal integer estimator [14]. The third step comes the key of fixed failure rate approach,identifying the relationship between failure rate and ratio test values. Verhagen proposed a root finding method to solve the problem [27], implementation of root finding method mainly use a Matlab function 'fzero'. More details are demonstrated to give a more intuitive description on the relationship in this section. The third step of previous procedure involves three sub-steps again:

1. Sample statistic. Dividing [0,1] into N small intervals, for example, $N = 1,000$. Statistics incorrectly fixed sample number $n_i, i = 1, 2, \ldots, N$ falling in each interval, calculating proportion of incorrectly fixed sample for each interval $p_i = n_i/N, i = 1, 2, \ldots, N$.
2. Finding relationship between failure rate and ratio test threshold. According to definition of failure rate (Seeing Eq. (34.15)), failure rate of ith interval can be numerically calculated by

$$p_{f,i} = \sum_{1}^{i} p_i, i = 1, 2, \ldots, N \tag{34.18}$$

   Then, the relationship between failure rate and ratio test threshold can be set up by replacing i in Eq. (34.18) by i/N.
3. Find μ for specific p_f according to the relationship. As the relationship is an irregular curve, it is difficult to find a simple model to fit the curve globally optimal. However, locally optimal is achievable, which means only 10–20 $p_{f,i}$ around specified p_f are picked out and fitting these $p_{f,i}$ by straight line or parabola. finally, μ can be determind by fitted curve.

An example to demonstrate the relationship between p_f and μ is given in Fig. 34.6. In this figure, 100,000 $\hat{a}$ samples are generated. The distribution of sample number and incorrectly fixed sample proportion respect to ratio test value are given. After Integral, the relationship between p_f and μ is given in the right sub figure. Figure 34.6 shows ratio test threshold varies with failure rate. Even for a very large ratio test value, saying close to 0, it is still possible to accept incorrect integer ambiguities.

One of the major shortcomings of the fixed-failure rate approach is the time efficiency. Performing 100,000 times LAMBDA to calculate one threshold is unacceptable for real-time positioning users. In order to meet real-time positioning requirement, all simulations can be performed in advance [23, 31]. Tabled the

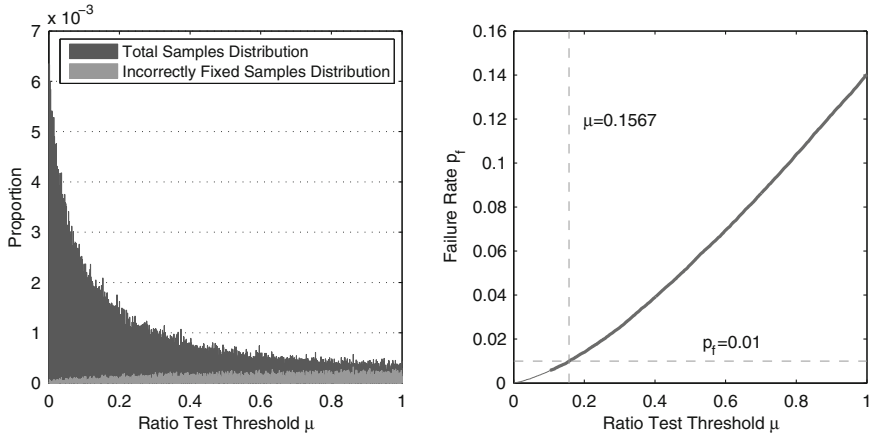


Fig. 34.6 An example of modelling relationship between ratio test value and failure rate, The *left* figure shows the distribution of total samples and incorrectly fixed samples respect to ratio test value. The *right* figure shows the relationship between p_f and μ, the relationship derived form integral of incorrectly fixed samples distribution. For this particular case, if $p_f = 0.01$, corresponding μ equals to 0.1567.

simulation results to show different integer least-squares failure rates against different ambiguity dimensions. Each ratio test threshold can be obtained by fixed-failure approach discussed above.

34.4 Simulation with Compass and GPS Constellation

As discussed above, the method of applying fixed failure rate approach is to prepare a table to establish the relationship between the integer least-squares failure rates and ambiguity dimensions in advance. In this section, simulation data for GPS and Compass constellations is presented.

Of the two key factors in ambiguity validation problem, hypothesis test with the fixed failure rate approach has been explained in previous section. The remaining issue is the ambiguity residual distribution. Ambiguity residual distribution can be characterised by $f_{\hat{a}}(x)$ and integer estimator. In this study, integer least-squares is chosen as integer estimator. But there are many influence factors in $\hat{a}$ distribution.

34.4.1 Function Model Dependency

$Q_{\hat{a}}$ is propagated from Q_y with design matrix in the function model. The function model's impact on $Q_{\hat{a}}$ is twofold: one is the number of real-valued parameters b in equation and another is the number of observations. For a short baseline, b only

involves 3 coordinate parameters. However, for longer distance baseline, additional tropospheric zenith delay and ionosphere delay parameters may be involved. If the observation number is determined, more parameters will definitely make the model weaker. Regarding the observation number, the simplest case is that only measurements from single-epoch are involved. A more complicate case is using Kalman filter that involves historical measurements and prior parameters variance information. According to generalized adjustment theory, historical information can be treated as virtual observations. With same parameters, more observations can improve model strength. It is impossible to enumerate all possible function models. In this context, we focus on the simplest case where 3 coordinate parameters together with ambiguity parameters are estimated from single epoch.

In order to make simulation as realistic as possible, GPS and Compass constellation is simulated to calculate the fixed failure rate ratio test threshold.

34.4.2 Stochastic Model Dependency

$Q_{\hat{a}}$ is propagated from Q_y, besides the propagation procedure, how to modelling Q_y is very crucial itself. Han improved RTK ambiguity resolution by improving stochastic model [3]. A popular prior stochastic model is $\sigma_\Phi = 3$ mm and $\sigma_P = 0.3$ m, correlation between code and phase, cross correlation between frequencies, temporal correlation are all ignored to keep it simple. Existing research shows there is correlation between observations, but the significance depends [26]. For simulation, It is impossible to consider all correlations,because it is difficult to quantify the correlation for general case. Unrealistic correlation model make the simulation worse than omittig correlations. In this study typical stochastic model is setting as $\sigma_\Phi = 10$ mm and $\sigma_P = 0.5$ m for conservative. Elevation dependency is ignored and all observations are assigned equal weight to keep function model as simple as possible. In order to demonstrate the influence of different stochastic model on fixed-failure rate ratio test threshold determination, three groups of experiments with different prior stochastic model is carried out and the result is shown in Fig. 34.7. In this study, 100,000 samples are simulated for each $Q_{\hat{a}}$, the fixed failure rate ratio test threshold μ are calculated for each $Q_{\hat{a}}$. The threshold μ are statistic against ambiguity dimension, mean and standard derviation of μ is presented in the figure.

According to the result, as observation become more precise, the threshold of fixed failure rate ratio test become larger, which means acceptance region of ratio test become lager, and correspondingly, probability of accepting $\check{a}$ is increased, vice verse. Another conclusion can be drawn from Fig. 34.7 is that fixed failure rate approach is not only sensitive to weight strategy,but also sensitive to unit weight factor σ_0 in Eq. (34.1). It means the prior observation precision must be precisely modelled for fixed failure rate approach. Either over evaluate or under evaluate the observation precision can influence the validation performance. However, σ_0 depends on receiver type, observation environment, satellite system et al. Hence, there is no universal stochastic model, the stochastic model need to be set up according to real situation.

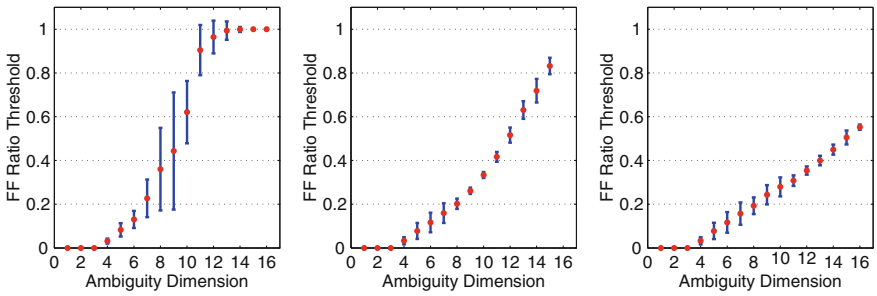


Fig. 34.7 Illustration of fixed failure rate ratio test threshold dependency on prior stochastic model for single frequency case. The prior observation precision is given as $\sigma_P = 0.25$ m, $\sigma_\Phi = 5$ mm (*left*), $\sigma_P = 0.5$ m, $\sigma_\Phi = 10$ mm (*middle*), $\sigma_P = 1$ m, $\sigma_\Phi = 20$ mm (*right*), 14,400 $Q_{\hat{a}}$ samples are used for each figure, including GPS, Compass, GPS+Compass scenarios, *Red* dot represent mean of ratio threshold and blue bar shows standard derivation of ratio threshold, Failure rate is given as $P_{f,ILS} = 0.001$.

34.4.3 Observation Number Dependency

As discussed, observation number can influence the model strength. Only regarding actual observation, the visible satellite number and frequency number determines observation number. Observation number depends on receiver capacity as well as location and time variation, receiver capacity refers to it is multi-frequency and multi-mode or not. In this study, two typical locations (around 30N and 30S) are chosen, and GPS and Compass satellite constellation are simulated. In order to make simulation as realistic as possible, real GPS and Compass constellation is generated. The Compass constellation follows [8], visible GPS satellite varies from 5 to 11 and visible compass satellite varies from 4 to 8. Dual constellation case is used to investigate high dimensional case. In order to reveal the relationship between fixed failure ratio test threshold and ambiguity dimension, several experiments results are merged into one figure. Another concern is the FF ratio test threshold dependency on frequency, triple frequency observation for both GPS and Compass are generated, the frequencies of Compass follows [8] as well. The relationship between frequency and ambiguity dimension is demonstrated in Fig. 34.8.

According to Fig. 34.8, higher dimensional ambiguity has a larger threshold for same failure rate. More frequency number can improve the ratio test threshold significantly. The results indicate that for dual-frequency case, If visible satellite number more than 9, the ratio test threshold almost equals to 1. It doesn't indicate the validation is superfluous in this scenarios, it means the statistical failure rate always smaller than given P_f in this case. Generally, more observation will result in larger threshold, and larger acceptance region.

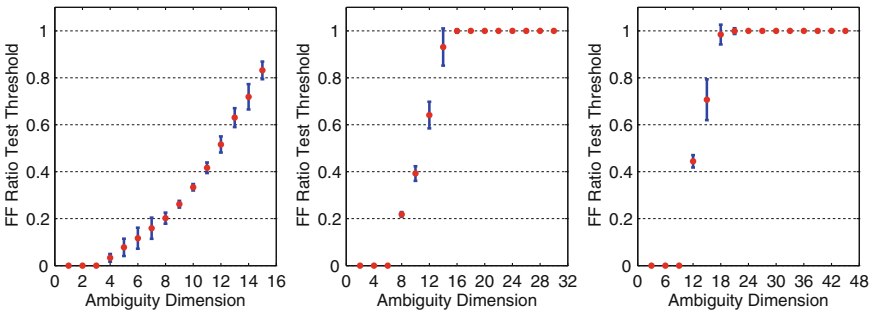


Fig. 34.8 Illustration of fixed failure rate ratio test threshold dependency on ambiguity dimension number and frequency number. The figure shows fixed failure rate ratio test threshold variation in single-frequency case (*left*), dual-frequency case (*middle*) and triple-frequency case (*right*). 14,400 $Q_{\hat{a}}$ samples are used for each figure, including GPS, Compass, GPS+Compass scenarios, Failure rate is given as $P_{f,ILS} = 0.001$. *Red dot* and *blue bar* show mean and standard deviation of fixed failure rate ratio threshold μ.

34.4.4 Failure Rate Dependency

The fixed failure rate ratio test threshold obviously depends on fixed failure rate. According to previous discussion, totally eliminating failure rate is almost impossible, and as a trade-off controlling type I error, a fixed failure rate should be chosen. The method to choose failure rate depends on the actual requirement.

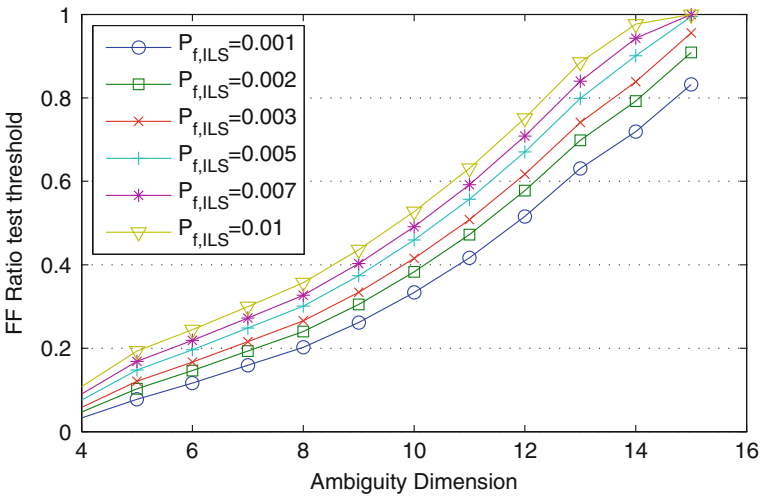


Fig. 34.9 Fixed failure rate ratio test threshold dependency on $P_{f,ILS}$ for single frequency case. The prior observation precision is given as $\sigma_P = 0.5$ m, $\sigma_\Phi = 10$ mm, 14,400 $Q_{\hat{a}}$ samples are used for each figure, including GPS, Compass, GPS+Compass scenarios, Different lines indicate ratio test threshold subject to different integer least-square failure rate. The markers indicate the mean of all μ samples.

Figure 34.9 shows the threshold for different failure rate. Generally, a larger failure rate means larger threshold, a larger acceptance region and more accepted $\check{a}$, but corresponding failure risk increased as well.

For real-time user, Fig. 34.9 can be transformed into a table, fixed failure rate ratio test threshold μ for different ambiguity dimension and different failure rate are all listed. With given ambiguity dimension number and failure rate, corresponding fixed failure rate ratio test threshold μ can be interpolated. User generated table is encouraged because user can give a more realistic stochastic model and function model according to real situation. The simulation result is more close to real world and the fixed failure rate can be a good way to control ambiguity estimation quality.

34.5 Concluding Remarks

This paper reviewed the ambiguity residual distribution first. Ambiguity residual distribution is not a normal distribution itself, but it can be expressed as an infinite sum of normal distributions. A general hypothesis test model for ambiguity validation problem was derived with ambiguity residual distribution. The general hypothesis test model explained the probability of type I error and type II error in ambiguity validation problem and the model is independent on any specific ambiguity validation method. In order to control type II error and ensure the reliability of ambiguity validation output, fixed failure rate approach is adopted. Ratio test is picked as an example to demonstrate the procedure of applying fixed failure rate approach. The key problem in fixed failure rate approach, the relationship between integer least-squares failure rate and ratio test threshold is systematically investigated. A full procedure of applying fixed failure rate ratio test is demonstrated.

Extensive fixed failure rate ratio test simulation based on GPS and Compass constellations is carried out. The simulation results indicate that ratio test threshold varies with function model, stochastic model, ambiguity dimension and failure rate. It is not reasonable to chose a fixed threshold for all senarios. Generally, a stronger model leads to a larger fixed failure rate threshold and larger acceptance region. However, the fixed failure rate approach is very sensitive to stochastic model. Hence, stochastic model should be as realistic as possible for ambiguity validation propose. Although fixed failure rate approach requires huge computation, the approach is still applicable for real-time users by generating a look-up table with given realistic function model and stochastic model. Once look-up table is generated, user can acquire fixed failure rate ratio test threshold by interpolating the table.

Acknowledgment This work was funded by cooperative research center for spatial information (CRCSI) project 1.01 'New carrier phase processing strategies for achieving precise and reliable multi-satellite, multi-frequency GNSS/RNSS positioning in Australia'.

References

1. Euler H, Schaffrin B (1991) On a measure for the discernibility between different ambiguity solutions in the static-kinematic GPS-mode. In: IAG, Symposium, pp 285–295
2. Frei E, Beutler G (1990) Rapid static positioning based on the fast ambiguity resolution approach fara: theory and first results. Manuscripta Geod 15(6):325–356
3. Han S (1997) Quality-control issues relating to instantaneous ambiguity resolution for real-time GPS kinematic positioning. J Geodesy 71(6):351–361
4. Hassibi A, Boyd S (1998) Integer parameter estimation in linear models with applications to GPS. Signal Proc, IEEE Trans on 46(11):2938–2952
5. Ji S, Chen W, Ding X, Chen Y, Zhao C, Hu C (2010) Ambiguity validation with combined ratio test and ellipsoidal integer aperture estimator. J Geodesy 84(10):597–604
6. Landau H, Euler H (1992) On-the-fly ambiguity resolution for precise differential positioning. In: Proceedings of ION GPS , pp 607–613
7. Li T, Wang J (2012) Some remarks on GNSS integer ambiguity validation methods. Surv Rev 44:230–238
8. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen P, Nakamura S (2012) Initial assessment of the compass/Beidou-2 regional navigation satellite system. GPS Solutions, 1–12
9. Neyman J, Pearson E (1933) On the problem of the most efficient tests of statistical hypotheses. Philosophical transactions of the royal society of London. Series A, Containing Papers of a Mathematical or Physical Character 231:289–337
10. Tabachnick B, Fidell L, Osterlind S (2001). Using multivariate statistics. Pearson Education, Inc
11. Tandy M, Young K (2012) Variable duration fixed failure rate ambiguity resolution. GPS Solutions, pp 1–7
12. Teunissen P (1995) The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. J Geodesy 70(1):65–82
13. Teunissen P (1998) On the integer normal distribution of the GPS ambiguities. Artif Satellites 33(2):49–64
14. Teunissen P (1999a) An optimality property of the integer least-squares estimator. J Geodesy 73(11):587–593
15. Teunissen P (1999b) The probability distribution of the GPS baseline for a class of integer ambiguity estimators. J Geodesy 73(5):275–284
16. Teunissen P (2002) The parameter distributions of the integer GPS model. J Geodesy 76(1):41–48
17. Teunissen P (2003a) A carrier phase ambiguity estimator with easy-to-evaluate fail rate. Artif Satellites 38(3):89–96
18. Teunissen P (2003b) Integer aperture GNSS ambiguity resolution. Artif Satellites 38(3):79–88
19. Teunissen P (2004) Penalized GNSS ambiguity resolution. J Geodesy 78(4):235–244
20. Teunissen P (2005a) GNSS ambiguity resolution with optimally controlled failure-rate. Artif Satellites 40(4):219–227
21. Teunissen P (2005b) Integer aperture bootstrapping: a new GNSS ambiguity estimator with controllable fail-rate. J Geodesy 79(6):389–397
22. Teunissen P (2005c) Integer aperture least-squares estimation. Artif Satellites 40:149–160
23. Teunissen P, Verhagen S (2009) The GNSS ambiguity ratio-test revisited: a better way of using it. Surv Rev 41(312):138–151
24. Teunissen P, Verhagen S (2011) integer aperture estimation a framework for GNSS ambiguity acceptance testing. Inside GNSS
25. Tiberius C, De Jonge P (1995) Fast positioning using the lambda method. In: Proceedings 4th international conference differential satellite systems, Citeseer, pp 1–8

26. Tiberius C, Jonkman N, Kenselaar F (1999) The stochastics of GPS observables. GPS World 10(2):49–54
27. Verhagen S (2005) The GNSS integer ambiguities: estimation and validation. PhD thesis
28. Verhagen S (2006) Improved performance of multi-carrier ambiguity resolution based on lambda method. In: Proceedings of Navitec 2006
29. Verhagen S, Teunissen P (2006a) New global navigation satellite system ambiguity resolution method compared to existing approaches. J Guidance Control Dyn 29(4):981–991
30. Verhagen S, Teunissen P (2006b) On the probability density function of the GNSS ambiguity residuals. GPS Solutions 10(1):21–28
31. Verhagen S, Teunissen P (2012) The ratio test for future GNSS ambiguity resolution. GPS Solutions, pp 1–14
32. Wang J, Stewart M, Tsakiri M (1998) A discrimination test procedure for ambiguity resolution on-the-fly. J Geodesy 72(11):644–653
33. Wei M, Schwarz K (1995) Fast ambiguity resolution using an integer nonlinear programming method. In: Proceedings of ION GPS 1995, pp 1101–1110
34. Xu P (2006) Voronoi cells, probabilistic bounds, and hypothesis testing in mixed integer linear models. Inf Theory, IEEE Trans on 52:7, 3122–3138

Chapter 35
Generalized-Positioning for Mixed-Frequency of Mixed-GNSS and Its Preliminary Applications

Shengfeng Gu, Chuang Shi, Yidong Lou, Yanming Feng and Maorong Ge

Abstract Modernized GPS and GLONASS, together with new GNSS systems, BeiDou and Galileo, offer code and phase ranging signals in three or more carriers. Traditionally, dual-frequency code and/or phase GPS measurements are linearly combined to eliminate effects of ionosphere delays in various positioning and analysis. This typical treatment method has imitations in processing signals at three or more frequencies from more than one system and can be hardly adapted itself to cope with the booming of various receivers with a broad variety of singles. In this contribution, a generalized-positioning model that the navigation system independent and the carrier number unrelated is promoted, which is suitable for both single- and multi-sites data processing. For the synchronization of different signals, uncalibrated signal delays (USD) are more generally defined to compensate the signal specific offsets in code and phase signals respectively. In addition, the ionospheric delays are included in the parameterization with an elaborate

S. Gu (✉) · C. Shi · Y. Lou
GNSS Research Center, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
e-mail: gsfjay@163.com

C. Shi
e-mail: shi@whu.edu.cn

Y. Lou
e-mail: ydlou@whu.edu.cn

C. Shi
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, 129 Luoyu Road, Wuhan 430079, China

Y. Feng
Faculty of Information Technology, Queensland University of Technology, 2 George Street, Brisbane, QLD Q4001, Australia
e-mail: y.feng@qut.edu.au

M. Ge
Helmholz Centre Potsdam, German Research Centre for Geosciences (GFZ), Telegrafenberg, 14473 Potsdam, Germany
e-mail: maor@gfz-potsdam.de

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_35,

consideration. Based on the analysis of the algebraic structures, this generalized-positioning model is further refined with a set of proper constrains to regularize the datum deficiency of the observation equation system. With this new model, uncalibrated signal delays (USD) and ionospheric delays are derived for both GPS and BeiDou with a large dada set. Numerical results demonstrate that, with a limited number of stations, the uncalibrated code delays (UCD) are determinate to a precision of about 0.1 ns for GPS and 0.4 ns for BeiDou signals, while the uncalibrated phase delays (UPD) for L1 and L2 are generated with 37 stations evenly distributed in China for GPS with a consistency of about 0.3 cycle. Extra experiments concerning the performance of this novel model in point positioning with mixed-frequencies of mixed-constellations is analyzed, in which the USD parameters are fixed with our generated values. The results are evaluated in terms of both positioning accuracy and convergence time.

Keywords Multi-GNSS · Multi-frequency · Generalized-positioning · Uncelebrated signal delay · Ionospheric delay

35.1 Introduction

The Chinese BeiDou navigation system at present consists of 16 satellites in orbits of which 14 are operational, independently providing preliminary navigation service in the Asia-Pacific Region. A full capability of regional service is expected in 2013. Users in the region can now take real benefit from multi-frequency observations from multi-GNSS constellations [1]. Concerning the performance of BeiDou signals, on board atomic frequency standards, as well as its contribution to the future GNSS systems, a wide range of valuable literatures has been published [2–5].

Of the 14 operational satellites, 10 satellites in active service since 2012 have already demonstrated that BeiDou can provide high precision relative positioning throughout the Asia-Pacific Region independently [6, 7]. Alongside the evolution of relative positioning technique, PPP, a stand-alone positioning technique, has also been widely developed as an efficient way to get high precision locations for both static and kinematic applications with global coverage [8, 9]. However, for the implementation of PPP algorithm, precise orbit and clock products must be available [10, 11]. Based on the BeiDou Experimental Tracking Stations (CETS) deployed by Wuhan University for a continuous monitoring of BeiDou Navigation Satellite System, it is now possible to estimate the BeiDou satellite orbit and clock at a repeatability of decimeter level with a modified version of the PANDA software [7, 12, 13].

In the PPP mode, typically, the linear combination known as ionosphere-free with dual-frequency observations is enabled within the GPS community [9, 10]. In processing BeiDou data, a similar combination of B1 and B2 is employed in the

preliminary studies [4, 7, 14]. However, since this linear combination is originally formulated for dual-frequency signals only, it does not necessarily the best solution for a broader variety of frequencies which is expected to be a big contribution of BeiDou system, providing that it is the first constellation with full consistency of three carriers ensured for all satellites [5]. Furthermore, PPP with low-cost, single-frequency receivers has also become a focus issue in recent years because of its huge potential applications [15–18]. Therefore, there is a need for a generalized-positioning algorithm which is navigation system independent and carrier number unrelated, in order to access the full capabilities of the multi-GNSS frequencies and signals, and it would also play a key factor for the expansion of BeiDou applications internationally.

The fundamentally new aspect of such a generalized-positioning technique is that all raw available signals from variety of frequencies of multi-GNSS are incorporated in a single parameter estimation system directly, which offers enormous flexibilities to the users equipped with whatever kind of receivers, along with numerous new challenges, such as the so-called uncalibrated signal delays (USD) as well as the ionospheric delays. The work by Schönemann et al. [19] was the first attempt that explicitly addressed the issues for future GNSS analysis with their focus on the problems that may be caused by such signal specific delays in both satellites and receivers. Shi et al. [18] develops a new formula with raw observations for PPP applications with an elaborate consideration of ionospheric parameterization methods, which has been demonstrated as an efficient way to describe ionospheric variations for precise positioning with both single- and dual-frequency receivers [18, 20]. However, there is little in-depth study of this generalized-positioning aspect with real tracking multi-frequency data from multi-GNSS receivers.

In this article, we began with the examination of the new observation formulae with a strong emphasis on the USD and ionospheric delays. The datum deficiency of the new estimation model is analyzed in detail and disposed with a proper set of constrains. Next, this generalized-positioning is demonstrated with the estimation of USD as well as ionospheric delays for both GPS and BeiDou constellations, and the generated products are evaluated in terms of repeatability and accuracy. Finally, the demonstration of this generalized model in standalone positioning with mixed-frequency of mixed-GNSS is carried out.

35.2 Generalized-Positioning Principle

Since the ionosphere-free combination cancels out the first-order ionospheric delays, it was almost regarded as a de-facto standard across the GNSS community for not only positioning, but also satellite orbit and clock determination. However, with numerous achievements in the ionospheric studies over the past few decades, the drawback of using ionosphere-free linear combination is gradually emerging in twofold: on the one hand, among three or more signals, how to define the optimal

combined observations is an issue. On the other, the more knowledge we gain about the ionosphere implies that more information that can be utilized in parameter estimation could be lost in ionosphere-free/ionosphere-reduced combinations in any form (including single differences and double differences, etc.). Therefore, the research trend in recent year is to directly use original observations instead of ionosphere-free combination.

The original observations of the GNSS pseudo-range and carrier-phase are generally expressed as Eq. (35.1) [19]

$$\left.\begin{aligned} P(sig,t)_r^s &= \rho_r^s(sig) + c\cdot(\sigma t(t)_r - \sigma t(t-\tau)^s) \\ &\quad + \sigma trop_r^s + ucd(sig)_r - ucd(sig)^s \\ &\quad + \sigma ion(sig)_r^s + \varepsilon_P(sig) \\ \Phi(sig,t)_r^s &= \rho_r^s(sig) + c\cdot(\sigma t(t)_r - \sigma t(t-\tau)^s) \\ &\quad + \sigma trop_r^s + upd(sig)_r - upd(sig)^s \\ &\quad - \sigma ion(sig)_r^s - \lambda(sig)\cdot N_r^s + \varepsilon_\Phi(sig) \end{aligned}\right\} \tag{35.1}$$

in which, $P(sig,t)_r^s$, $\Phi(sig,t)_r^s$ are pseudo-range and carrier-phase on signal *sig* for the specific satellite s and receiver r pair at epoch t in metric units, respectively; ρ is the geometric range that satellite and receiver vectors involved; c denotes the speed of light; σt stands for the clock error of receiver or satellite, while in the latter case, the signal travel time τ should be considered; the tropospheric delay is denoted as $\sigma trop$ as normal; while, N is by definition the integer phase ambiguity in cycle, together with the corresponding wavelength λ; finally, the remaining terms defined as uncalibrated code delay *ucd*, uncalibrated phase delay *upd*, as well as the ionospheric delay σion are specifically for the generalized model. For more details concerning this formula we refer to Schönemann et al. [19]. It is noted that there is an alteration for upd^s in our formulation: using the same symbol as that of σt^s but the negative sign since we believe that, usd^s should be regarded as a satellite clock-like parameter. In the following discussion, the units of the parameters may also vary, depending on the context.

35.2.1 Algebraic Structure of Generalized-Positioning Model

Let us assume that the system consists in j receivers (r_1, r_2, ..., r_j) and k satellites (s_1, s_2, ..., s_k), and suppose that all satellites are tracked by all receivers in m frequencies, then the linear observation system can be expressed in its matrix-vector form as follow:

$$\left.\begin{aligned} Y &= AX \\ var(Y) &= \Sigma \end{aligned}\right\} \tag{35.2}$$

in which $A \in \mathcal{M}_{n,p}$, denotes the range space of n-row and p-column, Σ is the variance of observations. In the context of generalized-positioning model, the following relations are held

$$\left.\begin{aligned} n &= j \cdot k \cdot m \cdot 2 \\ p &= (5 + 2 \cdot m) \cdot j + (4 + 2 \cdot m) \cdot k + (1 + m) \cdot j \cdot k \end{aligned}\right\} \tag{35.3}$$

where the p unknown parameters consists of

receiver specific parameters $(\mathbf{5 + 2 \cdot m})$: receiver coordinates, receiver clock error, zenith tropospheric delay, receiver uncalibrated code and phase delay on each frequency;

satellite specific parameters $(\mathbf{4 + 2 \cdot m})$: satellite coordinates, satellite clock error, satellite uncalibrated code and phase delay on each frequency;

line of sight specific parameters $(\mathbf{1 + m})$: ionospheric delays and ambiguity terms on each frequency.

To guarantee this system to be invertible, a necessary condition is that the number of observations should be not less than that of the parameters, i.e.

$$\begin{aligned} n - p = {} & j \cdot k \cdot m - 5 \cdot j - 2 \cdot m \cdot j \\ & - 4 \cdot k - 2 \cdot m \cdot k - j \cdot k \geq 0 \end{aligned} \tag{35.4}$$

suppose that all the sites are equipped with dual-frequency receivers, i.e. $m = 2$, relation (35.4) is then deduced to $n - p = j \cdot k - 9 \cdot j - 8 \cdot k \geq 0$, this condition (35.4) can be fully satisfied, given adequate receivers for satellite number $k \geq 10$. The statement can be further confirmed by the fact that, the observations are increased proportionally with the accumulation of epochs, while, the parameters are not due to the correlations for a set of parameters, e.g. $\sigma trop, \sigma ion, usd, N$, etc. between epochs.

As discussed above, one should not worried too much that $n \geq p$, thus, $rk(A)$ ($rk(\cdot)$ denotes the rank of matrix) is determined by its column rank r_D, while, the necessary and sufficient condition for a system to be invertible lies in the fact that $rk(A) = p$, i.e., A has the full column rank [21]. Nevertheless, as we shall see, a detailed investigation about matrix A suggests that $rk(A) = r_D = p - c < p$, which implies that this generalized-positioning system suffers from the datum deficiency.

By definition, the design matrix A can be divided into four parts $A_i (i = 1, 2, 3, 4)$, which correspond to the partial derivative with respect to the coordinates, clock errors and USDs, atmosphere delays and ambiguities, respectively. The matrix is presented below:

$$Y = AX = (A_1 \quad A_2 \quad A_3 \quad A_4) \cdot \begin{pmatrix} x_{coor} \\ x_{clk,usd} \\ x_{atmos} \\ x_{amb} \end{pmatrix} \tag{35.5}$$

A few notations are now defined for future reference:

$$z_s = (0 \quad 0 \quad \cdots \quad 0)^T \tag{35.6}$$

$$u_s = (1 \quad 1 \quad \cdots \quad 1)^T \tag{35.7}$$

$$Z_s = \begin{pmatrix} 0 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & 0 \end{pmatrix} \tag{35.8}$$

$$U_s = \begin{pmatrix} 1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & 1 \end{pmatrix} \tag{35.9}$$

as defined, z_s is a $s \cdot 1$ vector with zero entries and u_s is a $s \cdot 1$ vector with one entries, while, Z_s is a $s \cdot s$ matrix with zero entries and U_s is a $s \cdot s$ identity matrix.

Concerning the sub-matrix A_1, we express the GNSS observations in terms of coordinates, which implies a selection of coordinate reference frame that contains 7 degrees of freedom, i.e. 3 translations, 1 scale and 3 orientations. From an algebraic point of view, this reference frame selection leads to a datum disposition for A_1. Since geodetic observations are, in general, insensitive to some degrees of freedom, the coordinate system deficiency has been discussed on a wide range of publications, notably the work of Sillard and Boucher [22], in which a practical method is proposed to remove this datum deficiency without introducing unnecessary disturbing information.

35.2.2 Absolute Information Carried by Phase Observations

Concerning the case where observations are grouped in the order of receiver by receiver and the upper part is corresponding to code and the lower part to phase, the sub-matrix A_4 thus is written as

$$A_4 = \begin{pmatrix} Z_{k\cdot m} & 0 & \cdots & 0 \\ U_{k\cdot m} & 0 & \cdots & 0 \\ 0 & Z_{k\cdot m} & \cdots & 0 \\ 0 & U_{k\cdot m} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & Z_{k\cdot m} \\ 0 & 0 & \cdots & U_{k\cdot m} \end{pmatrix} \tag{35.10}$$

Considering the phase observations only, i.e., omitting the corresponding rows of pseudo-range in A_4, the reduced matrix $\bar{A}_4$ is the identical matrix $U_{j\cdot k\cdot m}$. As a result, all columns in the reduced matrices $\begin{pmatrix}\bar{A}_1 & \bar{A}_2 & \bar{A}_3\end{pmatrix}$ can be expressed as an linear combination of $\bar{A}_4$. Geometrically speaking, $\bar{A}_4$ is a set of the standard bases of the space $\mathcal{M}_{j\cdot k\cdot m}$, and any arbitrary vector $v \in \mathcal{M}_{j\cdot k\cdot m}$ is a linear combination of these bases. This implies that the phase observations of a single epoch do not contain any "absolute" information. Since the ambiguities are constant over a continuous tracking arc, the question is, can these parameters be separated from the ambiguities in an "absolute" sense? Actually, the answer depends on whether the corresponding coefficient vectors change apparently or not during the experimental time span.

To be specific, ω_t is defined as the coefficient vector of the interested parameter for epoch $t(t = t1, t2, \ldots, tn)$ and the linear relation is expressed as $\omega_t = \bar{A}_4 \cdot \mu_t$ with μ_t, which denotes the elements of ω_t with bases as $\bar{A}_4$, there are three following situations:

$\boldsymbol{\omega_t}$ **remains identical over time**, then μ_t keeps unchanged and the following relation holds true $\begin{pmatrix}\omega_{t1}^T & \ldots & \omega_{tn}^T\end{pmatrix}^T = \begin{pmatrix}\bar{A}_4 & \ldots & \bar{A}_4\end{pmatrix}^T \cdot \mu_t$ consequently, the corresponding parameter cannot be uniquely determined, actually, the main contribution of Φ to these kind of parameters is the improvement of precision, but hardly their "absolute" values, e.g. clock-like parameters σt and *usd*;

$\boldsymbol{\omega_t}$ **slowly vary with time**, the normal equation may be ill-conditioned, leading to unstable estimation of the specific parameter, i.e., coordinates within a short observation arc. This explains the limited positioning performance before convergence;

$\boldsymbol{\omega_{tn}}$ **rapidly vary with time**, it comes up with the unbiased estimation of the parameter even with carrier-phase only.

In the following analysis for A_2 and A_3, one may realize that there may no "absolute" information carried by phase observations even given a sufficiently long time span for certain parameters.

35.2.3 Clock and Uncalibrated Signal Delay

Since the regularization for A_2 reflects the definition of GNSS clocks indeed, it is recommended that one should recall the definition and realization of GPS time system and the parameter related to GPS clocks, i.e., differential group delay T_{GD} [23], which is also known as Differential Code Bias (DCB) in IGS working groups [24, 25].

Sub-matrix A_2 can be rewritten in detail as Eq. (35.11)

$$A_2 = \begin{pmatrix} e_{\sigma t_{r1}} & \cdots & 0 & e_{\sigma t^s} & e_{ucd_{r1}} & \cdots & 0 & e_{ucd^s} & e_{upd_{r1}} & \cdots & 0 & e_{upd^s} \\ 0 & \ddots & 0 & \vdots & 0 & \ddots & 0 & \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & e_{\sigma t_{rj}} & e_{\sigma t^s} & 0 & \cdots & e_{ucd_{rj}} & e_{ucd^s} & 0 & \cdots & e_{upd_{rj}} & e_{upd^s} \end{pmatrix} \tag{35.11}$$

In Eq. (35.11), e_x is the coefficient matrix of the corresponding parameter x, if we use $l_r = k \cdot m \cdot 2$ denotes the number of signals per receiver, together with (35.6–35.9), we obtain

$$e_{\sigma t_{ri}}(l_r, 1) = u_{l_r} \tag{35.12}$$

$$e_{\sigma t^s}(l_r, k) = \begin{pmatrix} u_m & 0 & \cdots & 0 \\ 0 & u_m & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & u_m \\ u_m & 0 & \cdots & 0 \\ 0 & u_m & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & u_m \end{pmatrix} \tag{35.13}$$

$$e_{ucd_{ri}}(l_r, m) = (\, U_m \quad U_m \quad \ldots \quad U_m \quad Z_m \quad Z_m \quad \ldots \quad Z_m \,)^T \tag{35.14}$$

$$e_{ucd^s}(l_r, (k \cdot m)) = \begin{matrix} U_{k \cdot m} & Z_{k \cdot m} \end{matrix}^T \tag{35.15}$$

$$e_{upd_{ri}}(l_r, m) = (\, Z_m \quad Z_m \quad \cdots \quad Z_m \quad U_m \quad U_m \quad \cdots \quad U_m \,)^T \tag{35.16}$$

$$e_{upd^s}(l_r, (k \cdot m)) = (\, Z_{k \cdot m} \quad U_{k \cdot m} \,)^T \tag{35.17}$$

in which, the coefficients $(l_r, \cdot)$ stand for the dimensions of each notation e_x.

Apart from the cumbersome form of A_2, the next issue is the linear dependence between σt_r and σt^s. Applying the addition operator to the corresponding column vectors of A_2, one can obtain

$$A_2^1 + A_2^2 + \cdots + A_2^j = A_2^{j+1} + A_2^{j+2} + \cdots + A_2^{j+k} = u_{j \cdot lr} \tag{35.18}$$

where A_2^i denotes the ith column of A_2. Similar dependence also exists between ucd_r and ucd^s, as well as upd_r and upd^s, since it is possible to shift ucd_r/upd_r by any constant and modify ucd^s/upd^s accordingly. As a result, additional information is needed for the separation between receiver-specific and satellite-specific delays σt, ucd and upd respectively.

Furthermore, recall that phase does not define the "absolute" values for these clock-like parameters, if we only consider the pseudo-range observations, i.e., omitting the corresponding rows of carrier-phase (lower part) in A_2, linear dependence can be also found between σt_r and ucd_r, as well as σt^s and ucd^s.

The linear dependence of the design matrix discussed so far in this section reveals the uncertainty of the time system underlying the GNSS technique itself. A solution of this generalized-positioning problem is the so-called free-network adjustment. In the remaining of this section, we turn the attention to the regularization of our specific model.

Numerically speaking, a singular problem can be solved by either generalized inverse [26] or regularization of the system with a set of constrains [27]. These two techniques are equivalent since each choice of generalized inverse corresponds to a choice of regularization constrains, and both lead to the identical result. In fact, such an algorithm leads to a solution class, among which, one solution can be converted to another theoretically through S-transformation [28]. However, as far as GNSS applications are concerned, this problem is more complicated: Firstly, a unique estimation is always essential in practical data processing; secondly, when subject to the GNSS service, e.g., International GNSS Service (IGS), it is mandatory that the datum definition are keep consistent on "server" and "client" side. As a result, extra conditions should be defined explicitly

$$\begin{cases} 0 = \sum_{i=1}^{k} \sigma\, t^{si}, \\ 0 = \sum_{i=1}^{k} ucd(f)^{si}, \ (f = 1, 2, \ldots, m) \\ 0 = \sum_{i=1}^{k} upd(f)^{si}, \ (f = 1, 2, \ldots, m) \end{cases} \tag{35.19}$$

Eq. (35.19) defines a proper set of conditions that consist of $(1 + 2m)$ equations to eliminate the dependence of σt, ucd and upd between receiver and satellite respectively. It is noted that the second and third relations are defined for each individual frequency (it is presume here that the pseudo-range or phase type is unique on each frequency, e.g., neglect the difference between C/A and P code). To separate clocks from UCD, following conditions are introduced

$$\begin{cases} 0 = ucd(1)_{ri}, & (i = 1, 2, \ldots, j) \\ 0 = ucd(1)^{si}, & (i = 1, 2, \ldots, k) \end{cases} \tag{35.20}$$

where the condition is applied to the "first" code signal for all involved receiver and satellite.

The particularly important aspect of the constrains (35.19) and (35.20) is the compatibility with the existing products, e.g., when it subject to GPS constellation, the condition (35.20) reads as $ucd(P1)_{ri} = 0$ and $ucd(P1)^{si} = 0$, together with $\sum_{i=1}^{k} ucd^{si} = 0$, the *ucd* on the other frequencies are actually given with respect to P1, which is exactly the current DCB products of IGS [24].

However, these constrains are still not adequate for the removal of the dependence between UPD and ambiguities. Concerning this deficiency, the constraints can never be added directly! On the one hand, by introducing artificial conditions on UPD, no one can guarantee the integer nature of ambiguities—the estimation of UPD depends on the estimation of "integer" ambiguities, on the other hand, without proper conditions on UPD, no integer property for ambiguities at all—the estimation of "integer" ambiguities depends on the estimation of UPD. A straightforward solution for this dilemma is to find the unbiased estimable parameters in a singular system [29]. Such a procedure is referred to as the double-difference (DD) operation in the GNSS community, which has a long evolutional path since the work of Dong and Bock [30] and Blewitt [31]. Instead of applying the DD operation on the observation level, De Jonge [32] discussed the DD GNSS processing strategy but on the equation level which revealed the potential of ambiguity resolution (AR) with un-difference (UD) observations, but still transform the original parameters to the unbiased estimable ones. It was until recently in the work of Ge et al. [33], the UPD are retrieved as products for the recovery of the integer nature in a single station PPP, independently, Laurichesse et al. [34] developed the concept of "integer" clock that enables AR-PPP.

In generalized-positioning, the DD operation is not considered neither on observation nor equation level, the intention is to solve the original parameters with raw observations. Although the parameters *upd* and *N* cannot be separated as a standard singular problem, it is still possible to append conditions on UPD in an iterative sense as enlightened from the approaches of Ge et al. [33] and Laurichesse et al. [34], i.e., first the constraints are set as $upd = 0$ and the UD ambiguity on individual signal is estimated with float value. After this, these float ambiguities are tested for stability and consistency to determine whether they are ready to be fixed or not. Once they are fixed, the fractional parts are extracted for re-adjustment to update *upd*, which are finally substituted as constrains in the generalized-positioning. In this way the constraints on *upd* may be added without destroy of the integer nature of *N*. However, in practice, the procedure is not a trivial task as described above, a reliable fix decision with UD ambiguity cause a huge challenge which still has a long way to go in the future indeed.

35.2.4 Atmospheric Delay

The formation of A_3 is subject to the understanding of the atmosphere. In our generalized-positioning model, the tropospheric delay over a single station is assumed to be isotropic and temporally correlated with a mapping function that converts the delay from zenith to slant directions. These assumptions are usually acceptable. Moreover, Boehm et al. [35] has introduced an example of mapping function and the temporal correlation can be applied as suggested by Tralli and Lichten [36]. These assumptions ensure the separability of tropospheric delay from other parameters. However, up to now, the ionospheric delay is regarded as an epoch-wise parameter for each satellite-receiver pair as presented in Eq. (35.1).

As proper constrains on ionosphere parameters are expected to improve the performance of the solution. Commonly accepted such functions are in forms of polynomial or trigonometric series [37, 38]. Unfortunately, these models do not describe the local TEC variations well. Consequently, Yuan and Ou [39] generalized the trigonometric serial functions by adaptive parameter set with some preliminary results. However, no further researches concerning the selection of parameters have followed. Considering the random behavior of ionosphere itself, both functional and stochastic parameters should be considered. Referring to the previous studies of involving authors in Shi et al. [18], some results are briefly summarized for convenience.

The fundamental step of our model is to refine the deterministic ionospheric delay over a station represented by a polynomial with a stochastic process for each satellite, thus the ionospheric delay $\sigma ion(sig)_r^s$ of each line of sight (LOS) for signal *sig* is written as

$$\sigma ion(sig)_r^s = \pm \bar{F}(f)_r^s \cdot (a_0 + a_1 dL + a_2 dL^2 + a_3 dB + a_4 dB^2 + r_r^s) \qquad (35.21)$$

where $\bar{F}(f)_r^s = F_r^s \cdot \alpha(f)$, F_r^s denotes the elevation-dependent mapping function and α stands for the factor which is frequency f related; a_0 is the average value of ionospheric delay over the station; a_1, a_2 and a_3, a_4 are the coefficients of the two second-order polynomials which are used to fit the horizontal gradients of east-west and south-north direction, respectively, and $a_i (i = 0, 1, 2, 3, 4)$ together describes the deterministic behavior of the ionospheric delay; while the scalar field r_r^s, including the small features superimposed on the polynomial, represents the stochastic component from a second-order stationary process; $dL(_r^s)$, $dB(_r^s)$ are the longitude and latitude difference between the ionospheric pierce point (IPP) and the approximate location of station, respectively. Through this decomposition into functional and random terms, more precise constraints are applied into the corresponding parameters.

The polynomial coefficients are estimated every 5 min, since vertical ionospheric delay is stationary during time lags below 10 min [40], while the temporal constraint of the ionospheric random parameter $r_r^s(t)$ as suggested by Shi et al. [18] are summarized in the following equations

$$C(h) = \begin{cases} c_s - c_s\{(3/2)(h/7200) - (1/2)(h/7200)^3\}, & 0 \le h < 7200; \\ 0, & h \ge 7200; \end{cases} \tag{35.22}$$

where the function $C(\cdot)$ is called a covariogram defined as $C(s_i - s_j) = cov(Z(s_i), Z(s_j))$ (s is the time variable, h is the time interval) [41], c_s corresponds to the constant when $r_r^s(t)$ becomes uncorrelated and should be a function of latitude B as

$$c_s(B) = \begin{cases} 0.042, & |B| < 5^\circ; \\ 00055 + 007\exp(-012|B|), & 5^\circ \le |B| < 50^\circ; \\ 0.005, & 50^\circ \le |B|; \end{cases} \tag{35.23}$$

Once Eq. (35.21) is adopted, the sub-matrix A_3 is determined by Eq. (35.24–35.26).

$$A_3 = \begin{pmatrix} e_{a_{r1}} & \cdots & 0 & e_{r_{r1}} & \cdots & 0 \\ 0 & \ddots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & e_{a_{rj}} & 0 & \cdots & e_{r_{rj}} \end{pmatrix} \tag{35.24}$$

$$e_{a_{ri}}(l_r, 5) = \begin{pmatrix} \bar{F}(f_1)_{ri}^{s1} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \cdots & \cdot & & & \cdots & & \\ \bar{F}(f_m)_{ri}^{s1} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \vdots & \cdot & & & \vdots & & \\ \bar{F}(f_1)_{ri}^{sk} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \cdots & \cdot & & & \cdots & & \\ \bar{F}(f_m)_{ri}^{sk} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ -\bar{F}(f_1)_{ri}^{s1} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \cdots & \cdot & & & \cdots & & \\ -\bar{F}(f_m)_{ri}^{s1} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \vdots & \cdot & & & \vdots & & \\ -\bar{F}(f_1)_{ri}^{sk} & \cdot & (1 & dL & dL^2 & dB & dB^2) \\ \cdots & \cdot & & & \cdots & & \\ -\bar{F}(f_m)_{ri}^{sk} & \cdot & (1 & dL & dL^2 & dB & dB^2) \end{pmatrix} \tag{35.25}$$

$$e_{r_{ri}}(l_r,k)=\begin{pmatrix}\bar{F}(f_1)_{ri}^{s1} & \cdot & (1 & \cdots & 0)\\ \cdots & \cdot & & \cdots & \\ \bar{F}(f_m)_{ri}^{s1} & \cdot & (1 & \cdots & 0)\\ \vdots & \cdot & & \vdots & \\ \bar{F}(f_1)_{ri}^{sk} & \cdot & (0 & \cdots & 1)\\ \cdots & \cdot & & \cdots & \\ \bar{F}(f_m)_{ri}^{sk} & \cdot & (0 & \cdots & 1)\\ -\bar{F}(f_1)_{ri}^{s1} & \cdot & (1 & \cdots & 0)\\ \cdots & \cdot & & \cdots & \\ -\bar{F}(f_m)_{ri}^{s1} & \cdot & (1 & \cdots & 0)\\ \vdots & \cdot & & \vdots & \\ -\bar{F}(f_1)_{ri}^{sk} & \cdot & (0 & \cdots & 1)\\ \cdots & \cdot & & \cdots & \\ -\bar{F}(f_m)_{ri}^{sk} & \cdot & (0 & \cdots & 1)\end{pmatrix} \tag{35.26}$$

it is noted that, the subscripts $^{s}_{r}$ accompany with dL and dB is omitted for simplification. The zero value of the non-diagonal elements in Eq. (35.24) implies that we do not impose any assumption on the correlation of ionosphere among different stations. This suggests that this model is suitable for the ionospheric delay estimation with a sparsely distributed GNSS network, or even a standalone receiver.

Due to the fact that, σt_r, σt^s, $\sigma trop$ and ucd_r are either frequency or satellite independent, and upd is constrained with zero or values extract from UD float ambiguity, these parameters can be separated from a_i without doubt. Furthermore, the random ionospheric parameter r_r^s is associated with a prior information as well as temporal correlation derived from Eqs. (35.22) and (35.23). The remaining terms to which special attention should be paid are a_i, ucd_s and N.

Since the coefficient vector of ucd_s is identical for all involved stations, suppose that $j > 1$, a_i is then estimable, benefiting from the different geographical and mapping factors for different stations. Concerning the single station generalized-positioning, the separability of a_i is based on the variation of its coefficient (geographical and mapping factor) along time. However, since a_i is only correlated within 5 min, the "absolute" ionospheric delay and UCD can hardly be estimated stably as only slow change is experienced for these coefficients. To improve the geometry of the design matrix, the ionospheric delays from an available model, e.g. the Global Ionospheric Map (GIM) are also considered as pseudo-observations with proper weight as suggested by Shi et al. [18].

Although one can impose more conditions on ionosphere, e.g., a longer time correlation assumption on a_i to improve the function condition number, the problem is that the exterior conditions may not follow the irregular spatial and temporal variations. In contrast, if the ionospheric delay is treated as an

epoch-wise parameter for each LOS as at the very beginning, there would be no mismodeling problem. However, σion turns out to be linear dependent with ucd_s and N. Therefore, the estimation of ionospheric products with GNSS observations is actually a tradeoff between the ionosphere model and its separability. It was only when the model attached can separate the ionospheric delay and ideally describe its behaviors, we can regard these solutions as unbiased estimations of the "absolute" ionospheric delays. Otherwise, the estimations are biased due to mismodeling or only "relative" values are obtained due to the datum deficiency.

35.2.5 Summary of Invertible Generalized-Positioning

With the in-depth analysis on the algebraic structure of generalized-positioning, the datum deficiency problem is reduced by either external conditions as introduced to clock and USDs or model constrains as imposed on ionospheric delays. Finally, together with model (35.2), we get the generalized-positioning system that is invertible

$$\left.\begin{array}{c} Y = AX \\ var(Y) = \Sigma \\ C = BX \end{array}\right\} \tag{35.27}$$

where, A is defined by Eqs. (35.5), (35.10), (35.11) and (35.24); B stands for the design matrix of external conditions as expressed by Eqs. (35.19) and (35.20); C denotes the constant vector which is filled with zero entries without AR, while with the corresponding items replaced with upd either from a server system or retrieved from float ambiguity for AR.

In practice, it is not always necessary to estimate all the parameters simultaneously, e.g., for the estimation of usd, σion, the coordinates for both satellites and receivers are assumed known, while for the positioning, those parameters usd can be constrained with priori information.

As demonstrated in the next section, this generalized-positioning concept is generally suitable for processing mixed-frequency, mixed-GNSS observations from a single or multi-receivers. It is a novel model system for real times estimation of clock, USD, ionospheric delays as well as coordinates.

Table 35.1 Datum description for the USD and ionosphere experiment

Product	Constellation	Time	Inter[sec]	Sites	Location
P2 *ucd*	GPS	001–031, 2010	30	20	Figure 35.1
P2 *ucd*	BeiDou	156–178, 2012	30	07	Figure 35.4
P3 *ucd*	BeiDou	149–151, 2012	30	03	Figure 35.6
upd	GPS	031, 2012	30	37	Figure 35.8
σion	GPS&BeiDou	156–178, 2012	30	07	Figure 35.4

35.3 Demonstration with USD and Ionosphere Estimation

35.3.1 Experimental Data Set

In order to validate the performance, this section adopts the generalized-positioning for USD determination, as well as ionospheric delay estimation with both GPS and BeiDou observations. The IGS final products including orbits and 30 s interval clocks are held fixed during the processing, while, for the BeiDou precise ephemeris, we refer to the solution of Shi et al. [7], the station coordinates are fixed ether from the IGS weekly combination or calculated with PANDA software. Five experiments are carried out in this section as listed in Table 35.1.

35.3.2 UCD Estimation for GPS

Presented in Fig. 35.2 is the average statistic over the 31 days for GPS UCD estimation on L2 frequency with different station set. It can be concluded that the results are promising for UCD estimations on the level of 0.1 ns with a sparse monitoring network. Comparison of these 4 strategies suggests that 20-sites gives the smallest STD of 0.088 ns, which has an improvement of about 16.3, 42.7 and 43.4 % with respect to 5-, 10- and 15-sites solution respectively. Furthermore, as discussed in Sect. 35.2.3, the constraints suggested by this work is consistent with the existing products when it subjects to GPS constellation only. In this case, ucd_{P2} is actually referred to as P1P2 DCB in the IGS community, thus it is possible to assess the accuracy of GPS *ucd* by comparison with the monthly DCB from Center for Orbit Determination in Europe (CODE) as shown in Fig. 35.3. It is further confirmed that there are no obvious bias exists for all these solutions with respect to the DCB values from CODE.

35.3.3 UCD Estimation for BeiDou

Different from the GPS system, since it is still under deployment, the BeiDou tracking data are only collected on a few stations with a limited coverage.

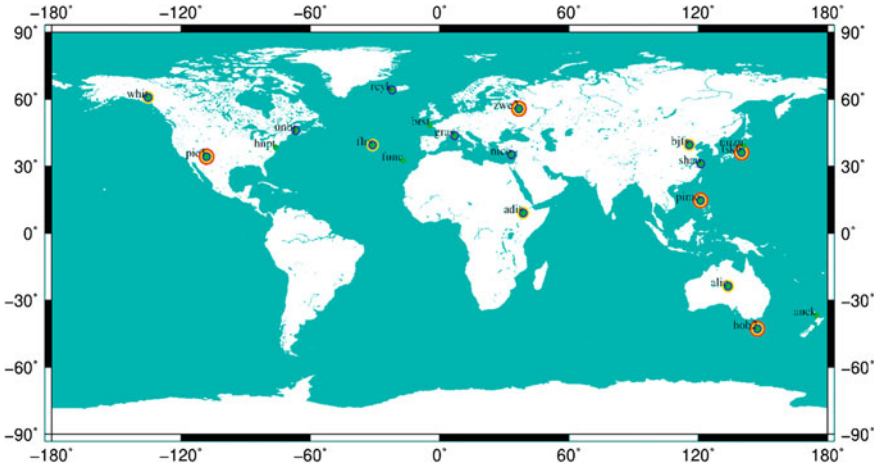


Fig. 35.1 Location of GPS stations for UCD estimation, four strategies are carried out with different set of stations involved that plotted in *green* (20), *blue* (15), *yellow* (10) and *red* (5), respectively

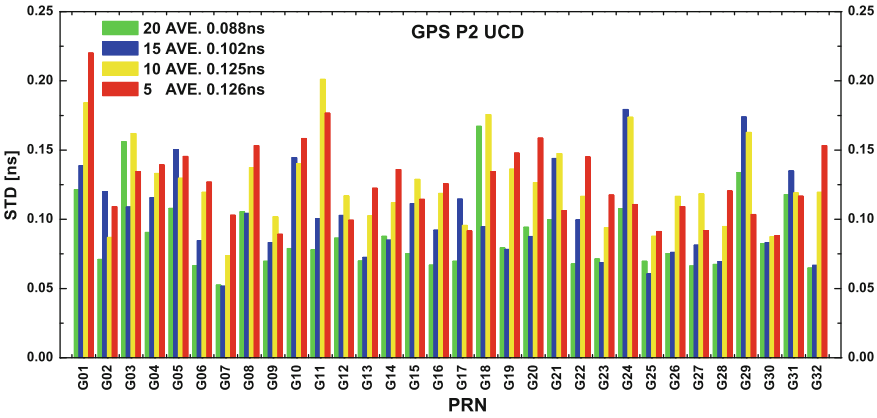


Fig. 35.2 STD of UCD estimation on L2 frequency for GPS over 001–031, 2010 with different set of stations as plotted in *green* (20), *blue* (15), *yellow* (10) and *red* (5), respectively

Experimental analysis in this context aims to assess the performance of BeiDou UCD estimation with a sparse network for both B2 and B3. Since no reference values available, the precision is evaluated in terms of repeatability. Regarding the B2 UCD estimation, observations from 7 stations as shown in Fig. 35.4 are collected over days 158–178, 2012 with a sample rate of 30 s, though these receivers are capable for the tracking of both GPS and BeiDou satellites, only the BeiDou date are processed in this experiment. Daily solutions of B2 UCD is plotted in Fig. 35.5, while the statistic is summarized in Table 35.2. As presented, the performance of B2 UCD is about 0.4 ns which is quite worse than that of GPS, this

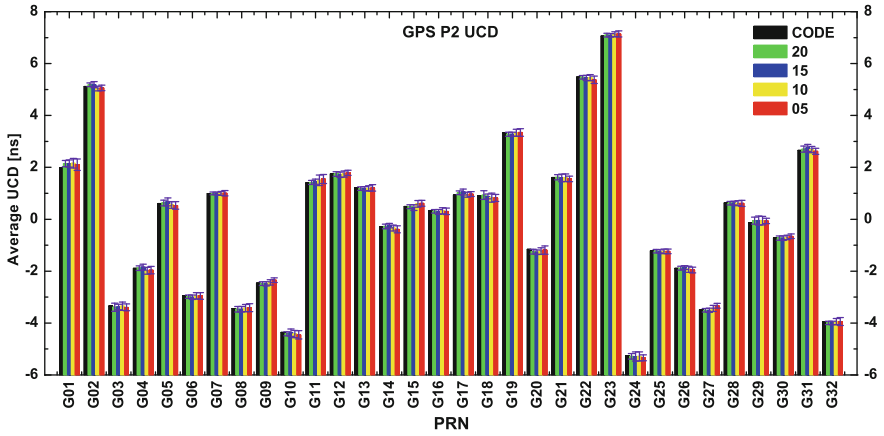


Fig. 35.3 Average UCD on L2 frequency for GPS over 001–031, 2010 with different set of stations as plotted in *green* (20), *blue* (15), *yellow* (10) and *red* (5), respectively, the monthly DCB values from CODE are plotted in *black* for reference, the error bar in purple shows the STD of the corresponding satellite

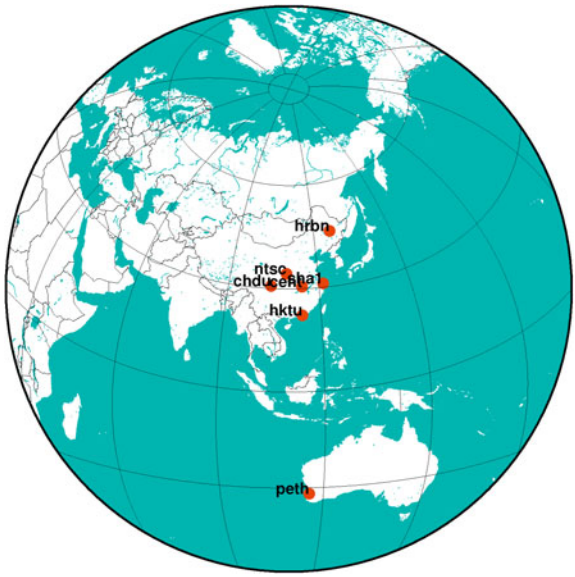


Fig. 35.4 Location of stations involved in BeiDou B2 UCD estimation and positioning

can be explained as a consequence of the imperfect quality of datum as well as the limited precision of associated products and models, e.g. orbit, clock and Phase Center Offset/Variance (PCO/PCV). Furthermore, the UCD on L2 frequency for IGSO(C06-C10) satellites has a precision of 0.254 ns, which is obviously better then that of GEO(C01-C05) satellites by a factor of 58 % due to the improvement of the orbit geometry. This also explains the higher precision of UCD estimation

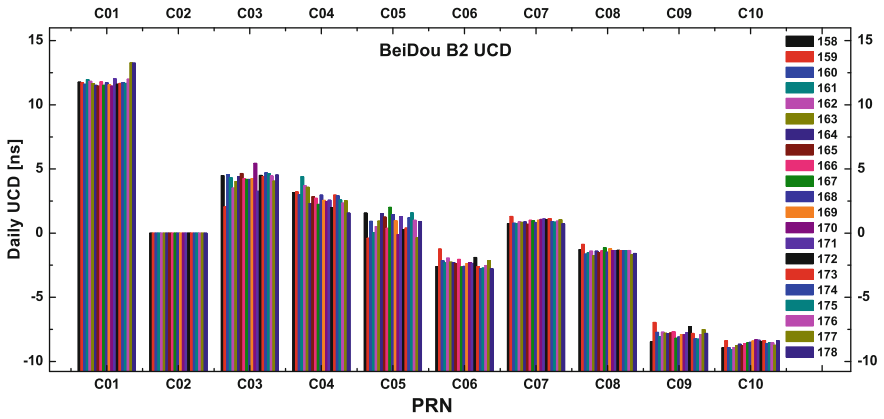


Fig. 35.5 Daily UCD on L2 frequency for BeiDou over 158–178, 2012

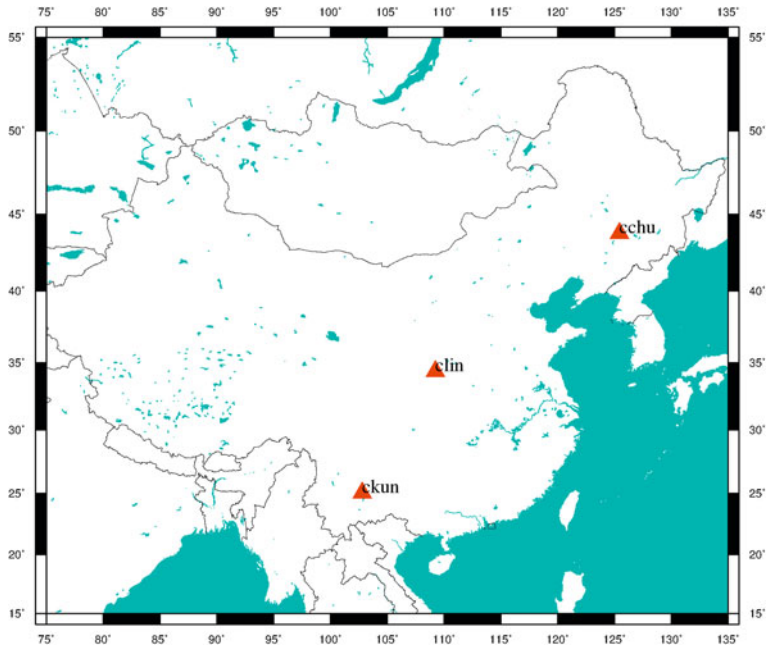


Fig. 35.6 Location of stations involved in BeiDou B3 UCD estimation

for GPS constellation which is consisted of MEO satellite only. Similarly, shown in Fig. 35.7 is the daily solutions of B3 UCD, since observations of only 3 stations as plotted in Fig. 35.6 over 3 days are available, the precision is rather limited at a level of 0.6 ns as calculated in Table 35.2, again, the STD improvement of UCD on L3 frequency for IGSO over GEO is significant, showing reduction from 1.07 to 0.31 ns.

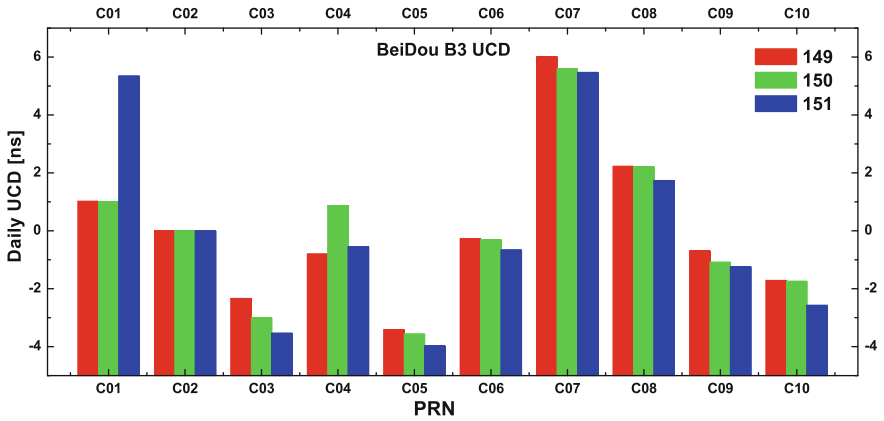


Fig. 35.7 Daily UCD on L3 frequency for BeiDou over 149–151, 2012

Table 35.2 Statistic of UCD parameters for BeiDou (ns)

SV	ucd_{B2}	$\sigma_{ucd_{B2}}$	ucd_{B3}	$\sigma_{ucd_{B3}}$
C01	11.85	0.496	2.458	2.499
C02	/	/	/	/
C03	4.23	0.663	−2.959	0.601
C04	2.79	0.616	−0.160	0.904
C05	0.82	0.679	−3.653	0.289
C06	−2.33	0.359	−0.417	0.215
C07	0.921	0.160	5.691	0.282
C08	−1.407	0.193	2.061	0.281
C09	−7.847	0.335	−1.01	0.282
C10	−8.614	0.224	−2.01	0.489
AVE	/	0.414	/	0.649

35.3.4 UPD Estimation for GPS

Due to the fact that the BeiDou data processing is still on its initial stage currently, only the UPD of GPS system are generated with the new model. Typically, the UPD estimation in UD date processing involves two steps: first the UD wide-lane ambiguities N_w are resolved for each receiver following the identification of the fractional part as UPD values, then the narrow-lane N_n is fixed with the help of N_w to extract the corresponding UPD. In this section we demonstrate the UPD estimation with generalized-positioning model, in which, the UPD on L1 and L2 frequencies are derived simultaneously. GPS data sets collected with 37 stations over China as shown in Fig. 35.8 are analyzed.

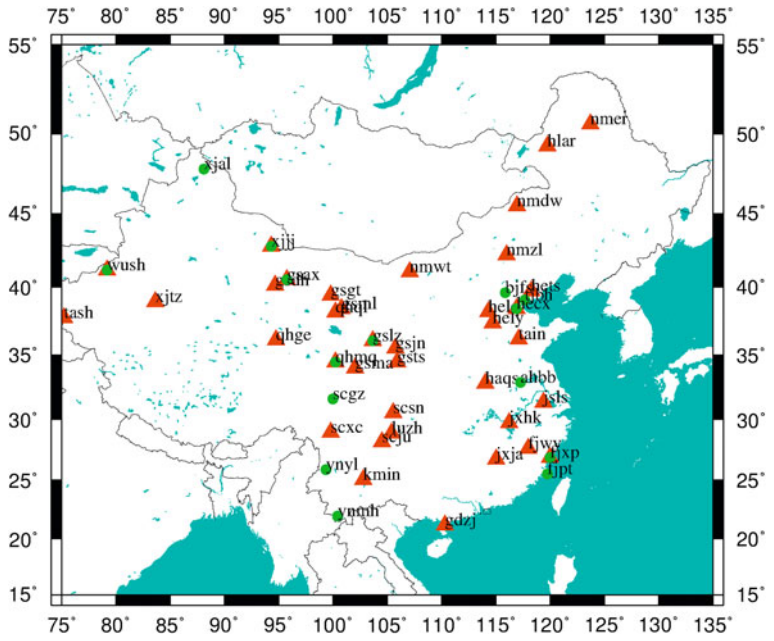


Fig. 35.8 Location of stations for the demonstrate of GPS UPD estimation with generalized-positioning model, the stations (37) denoted with *red triangle* are used for UPD estimation, while, the stations (16) denoted with *green circle* are used for positioning

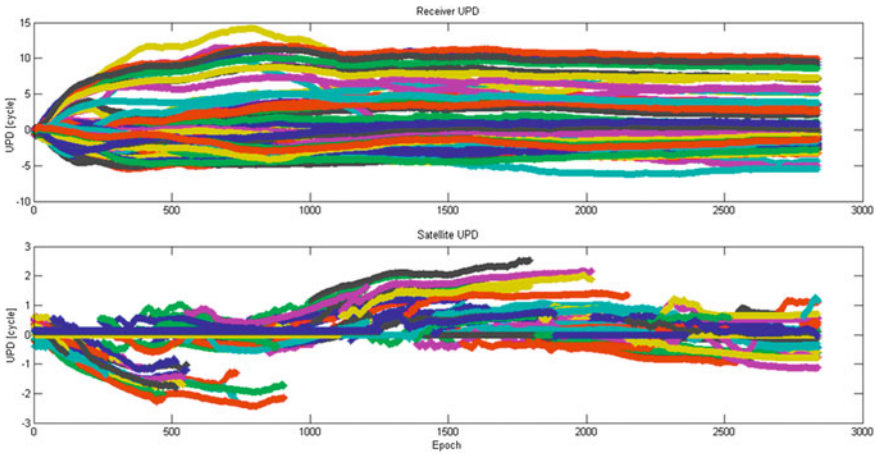


Fig. 35.9 *Upper panel* network filter UPD for receivers; *bottom panel* network filter UPD for satellites

Figure 35.9 illustrates the network filter UPD series for both receivers (upper panel) and satellites (bottom panel), it is clearly observed that the UPD on L1 and

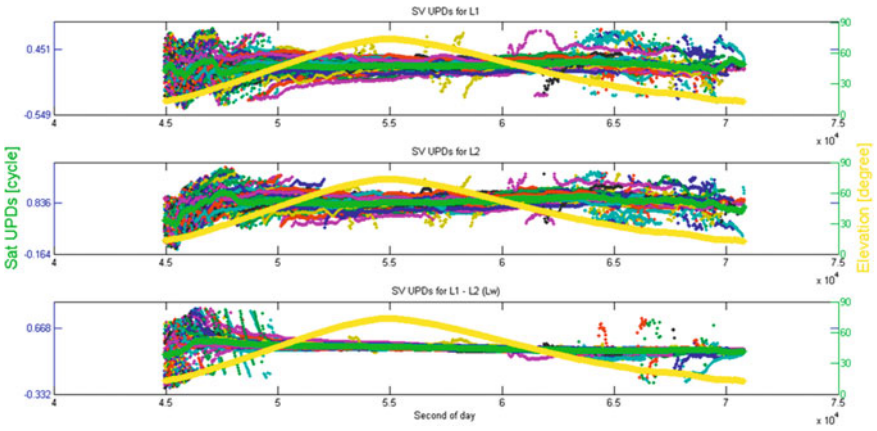


Fig. 35.10 *Upper panel* UPD series of PRN16 on L1 frequency estimated on each station; *middle panel* UPD series of PRN16 on L2 frequency estimated on each station; *bottom panel* UPD series of PRN16 on wide-lane retrieved with upd_{N1} and upd_{N2} on each station. The UPD of different stations are plotted with *different color*, while, the highlight *green series* is the network filter UPD series, the *yellow line* denotes the elevation angle

L2 frequencies can be hardly regarded as constant but has a slight fluctuation, while, the trend is consistent for different receivers as well as different satellites. Figure 35.10 compares the UPD series of PRN16 estimated on each receiver, the UPD on L1 and L2 frequencies plotted in the upper two panels suggested that the estimates from different receivers are agree with each other on a level of about 0.3 cycle for the satellites with a reasonable high elevation angle, i.e. $\geq 30°$, what's more, UPD of wide-lane upd_{NW} is retrieved with upd_{N1} and upd_{N2} and plotted in the bottom panel of Fig. 35.10. Obviously, the agreement of upd_{NW} is much better then that of upd_{N1} and upd_{N2} and more stable over time with a level of about 0.1 cycle.

35.3.5 Ionospheric Delay Estimation for GPS and BeiDou

As confirmed in Sect. 35.2.4, the generalized-positioning model is capable to derive the ionospheric effect. As an illustration, the observations of both GPS and BeiDou satellites collected on stations NTCS, CHDU and HKTU located at latitudes from 33.4° to 22.3° (see Fig. 35.4) on the day 168, 2012 are analyzed for the estimation of ionospheric delays. Shown in Fig. 35.11 is the zenith ionospheric delays based on the new method. Further in-depth evaluation on the performance of these delays is not presented in this work as it is beyond our discussion.

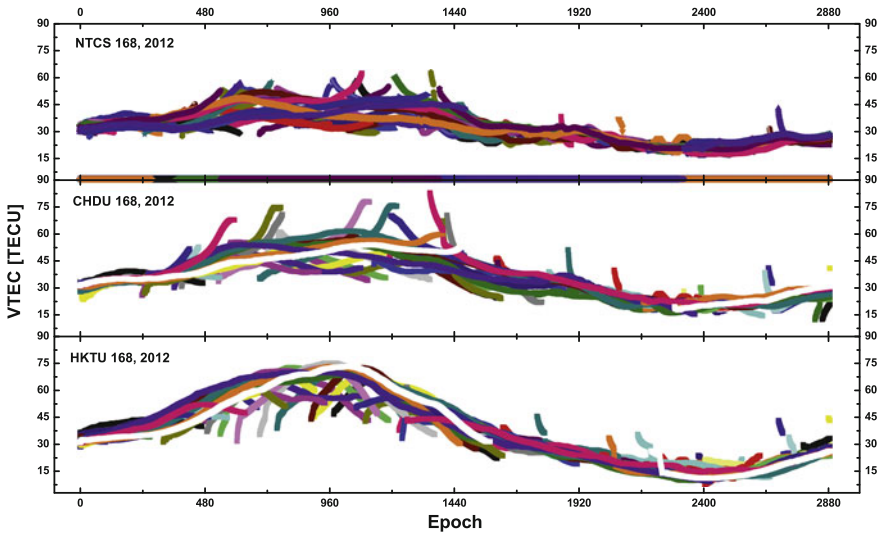


Fig. 35.11 Zenith ionospheric delay over NTCS (*upper panel*), CHDU (*middle panel*) and HKTU (*bottom panel*) on the day 168, 2012

35.4 Demonstration with Positioning for Mixed-Frequency of Mixed-GNSS

With the above network-based processing procedure, the UCD parameters of the satellites are held fixed in the standalone station positioning of this section. The precise ephemeris are again assumed known ether from IGS or estimated with PANDA software, and the ionospheric model values from GIM as pseudo-observations are added. The effectiveness of the generalized-positioning is assessed with different data set and different processing strategy to demonstrate its flexibility and accuracy. All of the experiments involved in the following sections are processed in simulated real-time, kinematic mode, receiver clocks are estimated epoch-by-epoch, station coordinates are estimated epoch-by-epoch with a constraint of $1\,\mathrm{m}/\sqrt{\mathrm{s}}$ between epochs. New ambiguity is inserted once a cycle slip is detected and observations identified as outliers are down-weighted.

35.4.1 Single Frequency Positioning for GPS and BeiDou

One of the main advantages of the new model is that the algorithm can be applied directly in data processing with mixed-frequency of mixed-GNSS, the test conducted in this section is to evaluated the single frequency PPP with GPS and/or BeiDou data that collected on stations as shown in Fig. 35.4. It is noted that only observations on L1 frequency are processed.

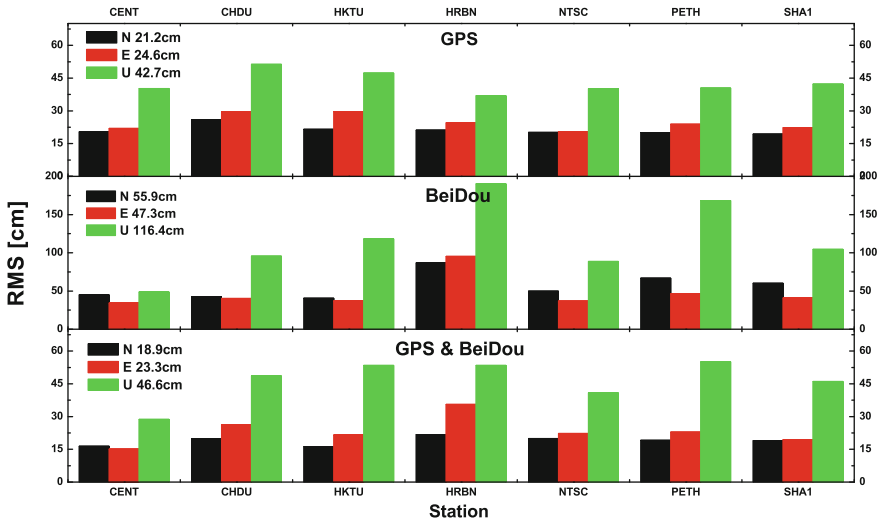


Fig. 35.12 Average RMS in the North, East an Up components over the days 158–178, 2012

Figure 35.12 is a plot of the average North-East-Up (NEU) RMS for GPS only (upper panel), BeiDou only (middle panel) and GPS and BeiDou (bottom panel) over the days 158–178, 2012. It is shown from Fig. 35.12 that the BeiDou only single frequency PPP has the largest RMS of about 0.5 m for horizontal and 1.15 m for vertical components, it is notable that the station HRBN has a RMS which is almost twice that of the other stations, this is mainly due to its high latitude location where the IGSO and GEO satellites have limited coverage. Comparison between the GPS only and GPS and BeiDou results, the contribution of the BeiDou to the GPS only positioning is insignificant, showing reduction from 21.2 and 24.6 cm to 18.9 and 23.3 cm for North and East components respectively, however, the vertical RMS increased from 42.7 to 46.6 cm. In addition, the positioning series of station CENT on the day 165, 2012 is plotted in Fig. 35.13 as an example for assessing their convergence, from which, it can be concluded that there is almost no visible convergence necessary for a sub-meter level positioning for all these three strategies.

35.4.2 Dual Frequency Positioning for BeiDou

The same data set as analyzed in the single frequency positioning is also enabled in this test but with dual frequency observations of BeiDou. The average daily NEU RMS values over the experiment days are illustrated in Fig. 35.14, where plotted in the upper panel is the results of forward filter, the bottom panel is the results of backward smooth. Concerning the filter results, the RMS of 10–20 cm in the

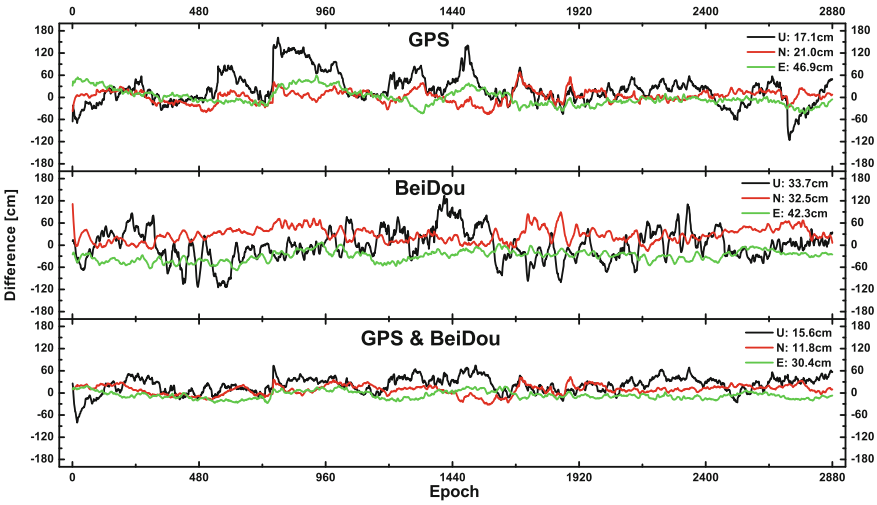


Fig. 35.13 Position differences in the North, East and Up components for station CENT on the day 165, 2012. The *top panel* shows the result with GPS only, while the *middle* is the result with BeiDou only, and the *bottom* for that of GPS and BeiDou

horizontal, and about 25 cm in the vertical components is obtained, which is unexpected worse than that of the smooth results which has an accuracy of 2 and 6 cm for horizontal and vertical components, respectively. Detailed analysis suggested that the BeiDou positioning takes a long convergence time to arrive at a stable positioning of centimeter level as illustrated in Fig. 35.15. For comparison, the filter result of GPS is also plotted in black, it is clearly observed that the BeiDou filter result take a much longer time for initialization than that of GPS. Recall that the separation between the coordinates and the ambiguities depends on the change of their coefficient vectors ω as discussed in Sect. 35.2.2, while the coefficient vector of MEO satellite vary quickly than that of GEO and IGSO, as a result, the convergence of GPS positioning accelerated significantly compared with BeiDou.

35.4.3 Ambiguity Resolution in GPS Point Positioning

With the *upd* values generated in Sect. 35.3.4 as constrain function $C = BX$ that substituted to Eq. (35.27), the ambiguity parameters N are integer defined, it is thus able to perform ambiguity resolution for point positioning. While, positioning with *upd* constraint as zero is also carried out which is refer to as float solution for the evaluation of fixed solution.

The positioning RMS of the 16 stations as shown in Fig. 35.8 with green circle is plotted in Fig. 35.16, it is noted that only the results after convergence (1 h later)

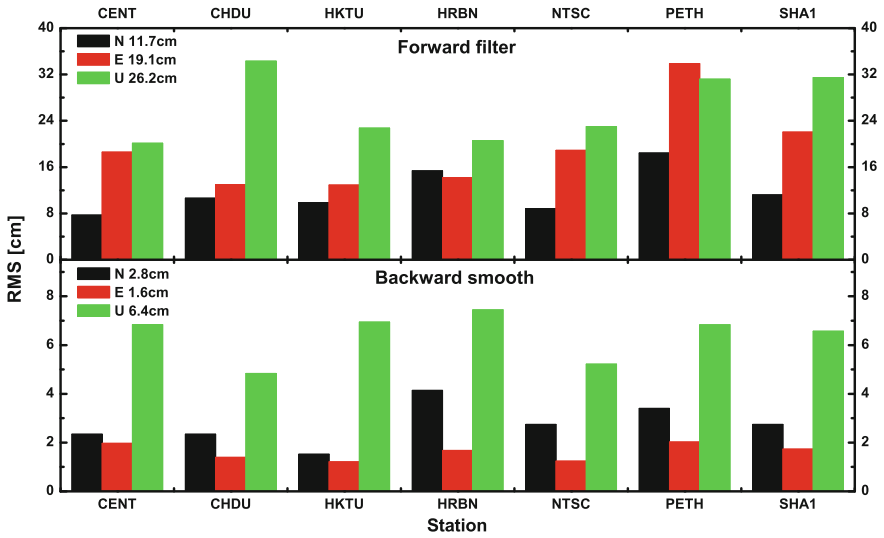


Fig. 35.14 Average RMS in the North, East an Up components over the days 158–178, 2012. The *top panel* shows the filter result, while the *bottom* shows the smooth result

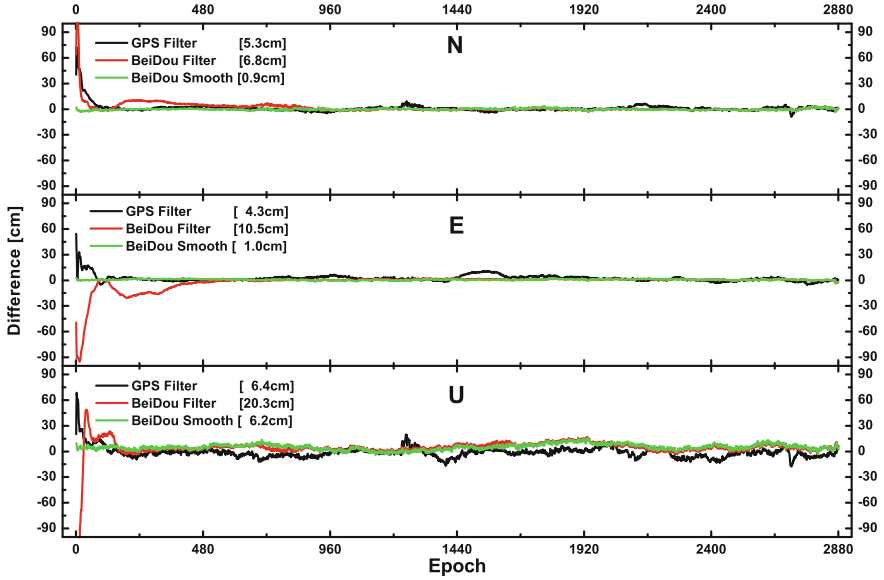


Fig. 35.15 Position differences in the North, East and Up components for station HKTU on the day 166, 2012. The serial plotted in *black* is the GPS filter result, while the *red series* denotes the result of BeiDou filter, and the *green series* denotes the result of BeiDou smooth

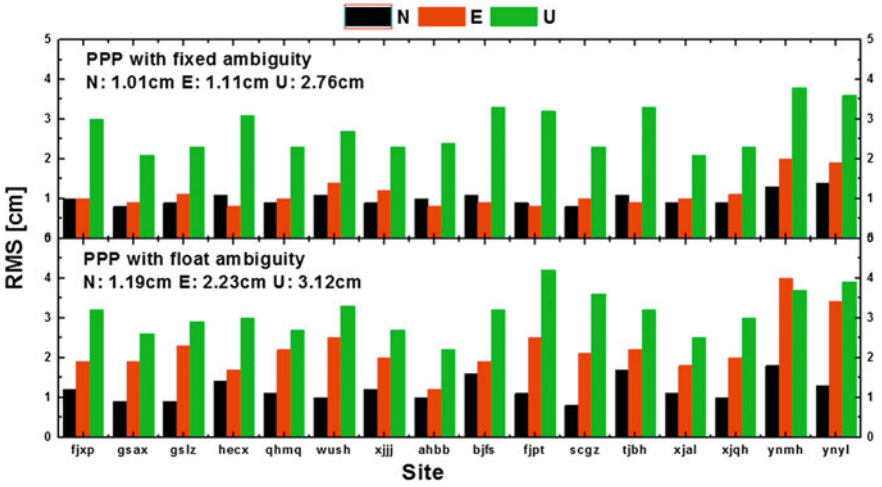


Fig. 35.16 RMS in the North, East an Up components on the days 031, only the results after convergence are considered in the statistic, 2012. The *top panel* shows the positioning result with integer ambiguity, while the *bottom* shows the positioning result with float ambiguity

are considered in the statistic. The contribution of AR to the accuracy of generalized-positioning is significant with a RMS reduction from 1.19, 2.23 and 3.12 cm to 1.01, 1.11 and 2.76 cm for North, East and Up components respectively. Concerning the performance of the initialization, the time series of station BJFS is presented in Fig. 35.17, from which, it is observed that though the series

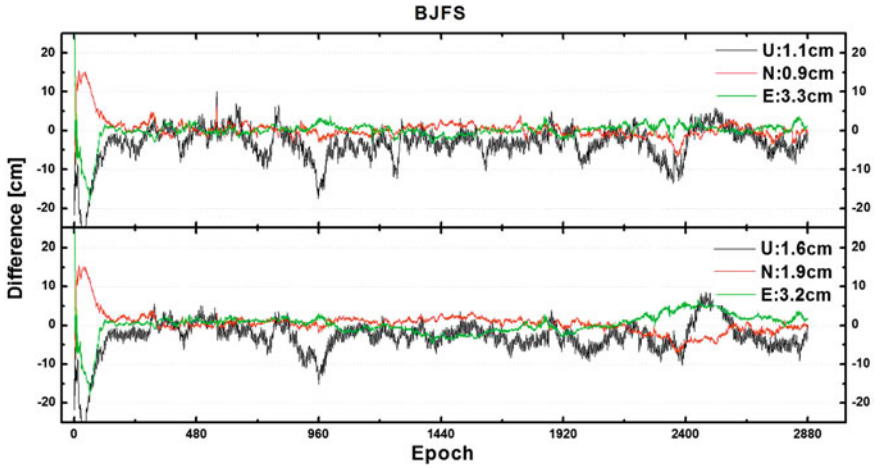


Fig. 35.17 Position differences in the North, East and Up components for station BJFS on the day 031, 2012. The *top panel* shows the positioning series with integer ambiguity, while the *bottom* shows the positioning series with float ambiguity

of horizontal components are more stable over time with ambiguity resolution, there is hardly any help for the acceleration of the convergence in our work.

35.5 Conclusions and Discussion

Obviously, under the multi-frequency and multi-GNSS environment, and taking into account the variety demands of users with a broad kind of receivers, an uniform algorithms which is navigation system independently, carrier number unrelated will be essential to access the full capabilities of the future navigation systems. This contribution proposed a generalized-positioning in which the raw observations of each individual signal are incorporated in a single parameter estimation system, furthermore, the signal special delays are introduced for the synchronization of different signals.

Based on the investigation of the algebraic structures of the generalized-positioning, its datum deficiency is disposed with a proper set of constrains. In addition, the *upd* constrains are suggested to applied in an iterative way, while the ionospheric effect over a station is fitted with a second-order polynomial and refined further by a stochastic process for each satellite.

One month of data collected in 2012 with different set of stations are processed with the new approach to assess the precision of UCD estimation of GPS, the results demonstrate that the accuracy of 0.1 ns are consistently achievable with about 10 stations. Concerning the BeiDou UCD, it is concluded that the precision is satellite related: for the IGSO satellites, the STD is about 0.25 and 0.31 ns for ucd_{B2} and ucd_{B3} respectively, while for the GEO satellites, the STD is about 0.61 and 1.07 ns for ucd_{B2} and ucd_{B3} respectively. Furthermore, the UPD parameters of GPS are generated with 37 stations all over China at a consistency of about 0.3 cycle for L1 and L2 frequencies.

Standalone positioning was also carried out for mixed-frequency of mixed-GNSS in both float and fixed solution. The single frequency test suggests the potential of sub-meter level positioning with BeiDou-only receivers with a accuracy of about 0.5 m for horizontal and 1.15 m for vertical, while, the GPS only positioning has a NEU RMS of 21.2, 24.6 and 42.7 cm respectively, and the contribution of BeiDou to the GPS positioning is limited in terms of accuracy since the favorable GPS environment in our case. Detailed analysis with dual-frequency BeiDou test illustrated that it is able to provide the centimeter level positioning as GPS, but need a much longer convergency time since only IGSO and GEO satellites are in active service currently. Further experiment of GPS positioning with integer ambiguities suggests that, the *upd* parameters can be applied as constrain functions easily in generalized-positioning for the ambiguity resolution to improve the accuracy.

Though we have demonstrated that the generalized-positioning is capable for the data processing of the current GNSS service, it is recommended that this concept is still on its initial stage with the following issues for future studies:

classification of the receivers: up to now, it is presumed that all receivers deliver the same type code or phase observation for the same frequency and satellite, which is in fact not the case, e.g., we have to distinguish between the P1/P2 and C1/P2 receivers, furthermore, the analysis of the correlation between these signal specific biases is dispensable;

constraints on UPD: since the formal method of free-network adjustment is not applicable for integer recovery of ambiguities, the constraints on UPD can be only added in an iterative manner which leads to a nonstandard processing strategy, and no one has proved that the coming result is the "best" estimation theoretically;

further development on the ionospheric model: the designed algorithm is suitable for standalone and multi-site data processing, however, the proposed ionospheric model disregard the correlation between stations which is surely a lost of information for the neighboring stations, there is no unified ionospheric model available for both single- and multi-sites data processing, what's more, the statistic analysis (see Shi et al. [18]) of the delays should be reprocessed under the solar-magnetic frame for a better solution;

extra data processing effort: accompany with the big flexibility of the novel algorithm is the increasing of computational power needed which is left for the future work.

References

1. Yang YX (2010) Prograss, contribution and challenges of compass/Beidou satellite navigation system. Acta Geodaetica et Cartographica Sinica 39(1):1
2. Hauschild A, Montenbruck O, Sleewaegen JM, Huisman L, Teunissen PJG (2012) Characterization of compass M-1 signals. GPS Solutions 16(1):117–126
3. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Riley S (2010) A COMPASS for Asia: first experience with the BeiDou-2 regional navigation system
4. Montenbruck O, Hauschild A, Steigenberger P, Hugentobler U, Teunissen PJG, Nakamura S (2012) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system. GPS Solutions 1–12. doi:10.1007/s10291-012-0272-x
5. Yang YX, Li JL, Xu JY (2011) Contribution of the Compass satellite navigation system to global PNT users. Chin Sci Bull 56(26):2813–2819. doi:10.1007/s11434-011-4627-4
6. Shi C, Zhao QL, Hu Z, Liu J (2012) Precise relative positioning using real tracking data from COMPASS GEO and IGSO satellites. GPS Solutions 1–17. doi:10.1007/s10291-012-0264-x
7. Shi C, Zhao Q, Li M, Tang W, Hu Z, Lou Y, Zhang H, Niu X, Liu J (2012) Precise orbit determination of Beidou satellites with precise positioning. SCIENCE CHINA Earth Sciences 55(7):1079–1086. doi:10.1007/s11430-012-4446-8
8. Bisnath S, Gao Y (2008) Current state of precise point positioning and future prospects and limitations. Observing our Changing Earth 133:615–623
9. Zumberge JF, Heflin MB, Jefferson DC, Watkins MM, Webb FH (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks. J Geophys Res 102(B3):5005–5017
10. Kouba J, Héroux P (2001) Precise point positioning using IGS orbit and clock products. GPS solutions 5(2):12–28

11. Shi C, Lou Y, Song W, Gu S, Geng C, Yi W, Liu Y (2011) A wide area real-time differential GPS prototype system in China and result analysis. Surv Rev 43(322):351–360. doi:10.1179/003962611X13055561708623
12. Liu J, Ge M (2003) PANDA software and its preliminary result of positioning and orbit determination. Wuhan Univ J Nat Sci 8(2):603–609. doi:10.1007/BF02899825
13. Shi C, Zhao Q, Geng J, Lou Y, Ge M, Liu J (2008) Recent development of PANDA software in GNSS data processing. In: Li D, Gong J, Wu H (eds) Proceedings of SPIE, the international society for optical engineering, SPIE, vol 7285, Bellingham, WA, p 72851S. doi:10.1117/12.816261
14. Steigenberger P, Hauschild A, Montenbruck O, Hugentobler U (2012) Performance analysis of compass orbit and clock determination and compass-only PPP. IGS Workshop
15. Gao Y, Zhang Y, Chen K (2006) Development of a real-time single-frequency precise point positioning system and test results. In: Proceedings of the 19th international technical meeting of the satellite division of the institute of navigation, pp 2297–2303
16. Le AQ, Tiberius C (2007) Single-frequency precise point positioning with optimal filtering. GPS Solutions 11(1):61–69. doi:10.1007/s10291-006-0033-9
17. Øvstedal O (2002) Absolute positioning with single-frequency GPS receivers. GPS Solutions 5(4):33–44. doi:10.1007/PL00012910
18. Shi C, Gu S, Lou Y, Ge M (2012) An improved approach to model ionospheric delays for single-frequency precise point positioning. Adv Space Res 49(12):1698–1708. doi:10.1016/j.asr.2012.03.016
19. Schönemann E, Becker M, Springer T (2011) A new approach for GNSS analysis in a multi-GNSS and multi-signal environment. J Geodetic Sci 1(3):204–214. doi:10.2478/v10156-010-0023-2
20. Li X, Ge M, Zhang H, Nischan T, Wickert J (2012) The GFZ real-time GNSS precise positioning service system and its adaption for COMPASS. Adv Space Res. doi:10.1016/j.asr.2012.06.025
21. Welsch W (1979) A review of the adjustment of free networks. Surv Rev 25(194):167–180
22. Sillard P, Boucher C (2001) A review of algebraic constraints in terrestrial reference frame datum definition. J Geodesy 75(2):63–73
23. ARINC Research Corporation (1997) Interfaca control document, Navstar GPS space segment/Navigation user interfaces, ICD-GPS-200, Revision C(IRN-200C-002)
24. Schaer S, Dach R (2010) Biases in GNSS analysis. In: IGS Workshop, Newcastle, England
25. Wilson B, Yinger C, Feess C, Shank C (1999) New and improved–the broadcast interfrequency biases. GPS World 10(9):56–66
26. Rao CR, Mitra SK (1971) Generalized inverse of a matrix and its applications. Wiley, New York
27. Blaha G (1982) Free networks: minimum norm solution as obtained by the inner adjustment constraint method. J Geodesy 56(3):209–219
28. Baarda W (1981) S-transformations and criterion matrices. Netherlands Geodetic Commission, Dalft
29. Grafarend E, Schaffrin B (1974) Unbiased free net adjustment. Surv Rev 22(171):200–218
30. Dong DN, Bock Y (1989) Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in California. J Geophys Res 94(B4):3949–3966
31. Blewitt G (1989) Carrier phase ambiguity resolution for the global positioning system applied to geodetic baselines up to 2000 km. J Geophys Res 94(B8):10187–10203
32. De Jonge PJ (1998) A processing strategy of the application of the GPS network. Delft University of Technology. Netherlands Geodetic Commission, Delft
33. Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. J Geodesy 82(7):389–399

34. Laurichesse D, Mercier F, Berthias JP, Broca P, Cerri L (2009) Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP and satellite precise orbit determination. Navigation 56(2):135–149
35. Boehm J, Niell A, Tregoning P, Schuh H (2006) Global mapping function (GMF): a new empirical mapping function based on numerical weather model data. Geophys Res Lett 33(7):L07304. doi:10.1029/2005GL025546
36. Tralli DM, Lichten SM (1990) Stochastic estimation of tropospheric path delays in global positioning system geodetic measurements. J Geodesy 64(2):127–159. doi:10.1007/BF02520642
37. Mannucci AJ, Wilson BD, Yuan DN, Ho CH, Lindqwister UJ, Runge TF (1998) A global mapping technique for GPS-derived ionospheric total electron content measurements. Radio Science 33(3):565–582. doi:10.1029/97RS02707
38. Schaer S (1999) Mapping and predicting the Earth's ionosphere using the global positioning system. Ph.D. Thesis, University of Bern, Switzerland
39. Yuan Y, Ou J (2004) A generalized trigonometric series function model for determining ionospheric delay. Prog Nat Sci 14(11):1010–1014
40. Blanch J (2003) Using Kriging to bound satellite ranging errors due to the ionosphere. Ph.D. Thesis, Stanford University, U.S.A.
41. Cressie N (1993) Statistics for spatial data. Wiley, New York, p 928

Chapter 36
Preliminary Study on the Variation of Zenith Tropospheric Delay with Altitude

Jingyang Zhao, Shuli Song and Wenyao Zhu

Abstract Tropospheric delay is a major source of error for ground-to-air, air-to-ground observation and various types of radio navigation positioning technology, and it is usually corrected by tropospheric delay models for the majority of real-time navigation users. Traditional tropospheric delay models (such as Hopfield, Saastamonion, UNB3, etc.) take the cumulative delay of the whole atmosphere into account, and surface meteorological data is needed as the input to some of them. It's hard to meet the requirement of the atmospheric delay correction for the real-time user at arbitrary time and height by using the old models. We consequently adopt the method modeling the zenith tropospheric delay (ZTD) with the altitude directly to solve this problem. In this paper, the ZTDs at different altitude were calculated by the ERA-Interim reanalysis data in 2004 at 1.5° grid points, and the spatiotemporal characteristics of ZTD varying with altitude was studied detailedly. A fitting analysis between ZTD and altitude was carried out by a piecewise function afterwards. The accuracy of the fitting function above was assessed carefully, and we also tested it by the ZTDs come from 45 IGS (International GNSS Service) stations distributed all over the world from 2008 to 2010. The results are as following: The deviation of the fitting and the integral ZTDs come from ERA-Interim is less than 1 cm at all the grid point, and which is stable in the range of 45 km high. The fitting and IGS ZTDs comparisons at 45 stations show a mean yearly difference of no more than 1 cm with a mean SD of about 4.3 cm. The piecewise fitting function we constructed could be applied as a preliminary model of the ZTD varying with altitude, and has important application value in real-time navigation and positioning.

Keywords Zenith tropospheric delay · Altitude · ERA-Interim

J. Zhao (✉) · S. Song · W. Zhu
Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai, China
e-mail: jyzhao7@gmail.com

J. Zhao
University of Chinese Academy of Science, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_36,

36.1 Introduction

Tropospheric delay is a major source of error for ground-to-air, air-to-ground observation and various types of radio navigation positioning technology. For example in the Global Navigation Satellite System (GNSS), the impact of the tropospheric delay can reach 2 m in the zenith direction and may be more than 20 m when the signal propagating direction deviates from the zenith towards the horizon direction. The delays of the signal propagating path usually calculated by the mapping function after the zenith tropospheric delay (ZTD) is obtained for the real-time navigation and positioning users. Thus, it's very important for the ZTD to be estimated exactly.

The real-time ZTD is often calculated by ZTD models, and Hopfield, Saastamoinen, EGNOS, UNB3 model are the most commonly used models at present. The two traditional tropospheric models, Hopfield and Saastamoinen model, can estimate ZTD at decimeter to centimeter precision level, but the real-time meteorological parameters must be obtained as their input. It's obvious that they are not perfect chooses for the real-time users. Collins and Langley [1] introduced UNB3 model which is based on the Saastamoinen model. In this model, the average and the annual changes amplitude of five meteorological parameters (pressure, temperature, water vapor pressure, temperature lapse rate and water vapor pressure decline rate) derived from the U.S. standard atmosphere data in 1966 are provided on the mean sea level for every 15° latitude band. The UNB3 model has been developing for the last several years [2–4]. Similar to the UNB3 model, the European Union established the EGNOS model, which is based on the 1° grid European Centre for Medium-Range Weather Forecast (ECMWF) data. Known as the model recommended by the International Civil Aviation Organization, the EGNOS model is widely used in the Satellite-Based Augmentation Systems of the USA, European, and Japan [5–7].

All the models mentioned above take the cumulative delay of the whole atmosphere into account, and there are two normal methods when the delay correction at arbitrary altitude is required. The first method is to interpolate the meteorological parameters to the corresponding height and then calculate the ZTD by the Hopfield or Saastamoinen model. However, besides the difficulty of obtaining real-time meteorological parameters, the calculation of the parameter interpolation is quite complicated and liable to great accumulative error. The other way is to compute it by EGNOS or UNB3 model, but the resolution of the parameters they offer is insufficient, so they only reflect the global tropospheric profile while lack of the details of the regional features. In our previous study, a new model called SHAO was established based on the NCEP and ECMWF data, including SHAO-C where 'C' stands for China and SHAO-G where 'G' stands for global. SHAO-C model is more accurate than EGNOS model for users in China [8] while a comparison between SHAO-G model and EGNOS model also showed that the former is more reliable [9]. We modeled the ZTD with altitude directly in the SHAO model so that the calculation of the delay becomes simple, but only the

variability of ZTD with altitude below 10 km is given in this model and the shortcoming of using quadratic polynomial to model the relationship between ZTD and altitude is founded. Consequently, it's necessary for us to study the spatio-temporal characteristics of ZTD with respect to altitude so as to extend the applicability of the SHAO model.

In this paper, the ZTDs at different elevations were computed by the ERA-Interim reanalysis data in 2004 at 1.5° grid points, and their variability along the height at different location and different time was studied detailedly. Then, a fitting analysis between ZTD and altitude was carried out by a piecewise function. The accuracy of the fitting function above was assessed carefully afterwards, and we also tested it by the ZTDs come from 45 IGS (International GNSS Service) stations distributed all over the world from 2008 to 2010.

36.2 Data

ERA-Interim is the latest global atmospheric reanalysis produced by ECMWF. Special emphasis is placed on various difficulties encountered in the previous production, including the representation of the hydrological cycle, the quality of the stratospheric circulation, and the consistency in time of the reanalyzed fields [10]. We mainly used the upper-air data of ERA-Interim, whose time resolution is 6 h, namely at 0, 6, 12, 18 UTC, and horizontal resolution is 1.5° × 1.5°. Each of the grid point data has 37 pressure levels and reaches about 45 km at the top of the level. The meteorological data we used are pressure, altitude, temperature, specific humidity and relative humidity.

The ZTD comes from IGS, as known to all, is the most accurate tropospheric delay in the world, whose precision is roughly 4 mm. In order to verify the accuracy of the ZTD calculated by the fitting function, ZTD time series from 2008 to 2010 of 45 IGS stations were used in this research. The distribution of the 45 stations is shown in Fig. 36.1.

36.3 Data Analysis

According to the study of Chen et al. [11, 12], the accuracy of the ZTD derived from ECMWF data is about 3 cm, which meets the requirement of our study. To get a ZTD series at different altitude, the ZTD was estimated at every pressure level of the EAR-Interim data using the integration method, the formulas are as following:

$$e = h \times P/0.622 \tag{36.1}$$

$$N = k_1(P - e)/T + k_2 \times e/T + k_3 \times e/T^2 \tag{36.2}$$

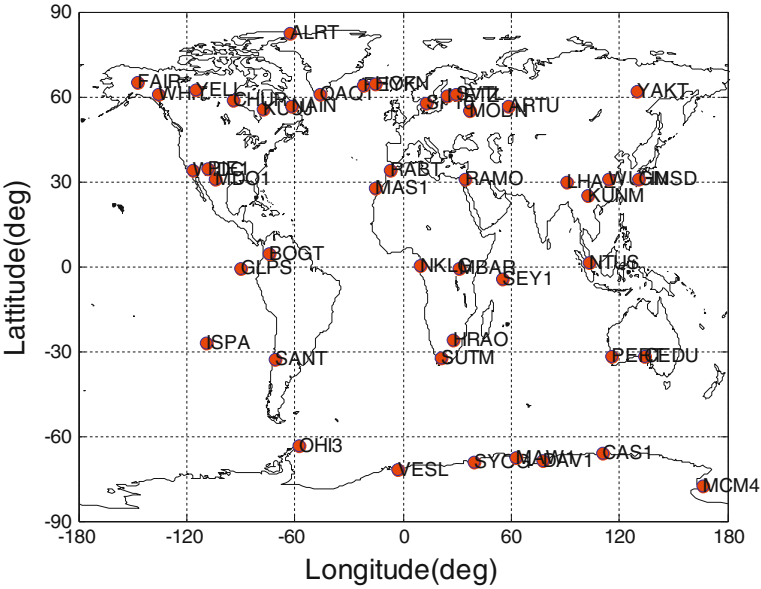


Fig. 36.1 Distribution of IGS stations

where N is the total refraction, P is the atmospheric pressure (hpa), e is the vapor pressure (hpa), h is the specific humidity, and k1 = 77.604 K/P, k2 = 64.79 K/Pa, k3 = 337600 K^2/Pa. After calculating the total refraction, the ZTD was derived by the formula:

$$\begin{aligned} \mathrm{ztd}_i &= 10^{-6} \int_{H_i}^{\infty} N \cdot dH \\ &= 10^{-6} \sum_{n=2}^{i} \frac{N^{n-1} + N^n}{2} \cdot \Delta H_{n-1,n} \end{aligned} \tag{36.3}$$

where N^n is the total refraction at the nth level, $\Delta H_{n-1,n}$ is the thickness between the different level, i = 2, 3,…,37, n = 1, 3,…,37.

In order to minimize the error in ECMWF-derived ZTD, the ZTD above the top level need to be added after the integration method. Because there is no meteorological data above the top level, the Saastamoinen model was used, and meteorological parameters of the top level were used as the input of the model. The formulas are as below:

$$e = rh \times 6.11 \times 10^{\frac{7.5T}{T+273.15}} \tag{36.4}$$

$$f(\varphi, H) = 1 - 0.00266 \cos 2\varphi - 0.00028H \tag{36.5}$$

$$ZTD\_S = 0.002277 \times \frac{\left[P + \left(0.05 + \frac{1255}{T+273.15}\right)e\right]}{f(\varphi, H)} \tag{36.6}$$

where rh is the relative humidity, ZTD_S is the Saasamoinen-derived ZTD. Finally the total ZTD at every level are given by:

$$\mathrm{ZTD_i}(\mathrm{H_i}) = \mathrm{ztd_i} + \mathrm{ZTD\_S} \tag{36.7}$$

where $\mathrm{H_i}$ is the corresponding height of the ith level.

36.4 Analysis of the Variability of ZTD with Altitude

36.4.1 Characteristics of ZTD with Altitude

Figure 36.2 shows the variation of ZTD with increasing height on the basis of evenly distributed latitude and longitude at 12:00 UTC, January 16, 2004. As can be seen clearly in Fig. 36.2, the ZTD value decreases with increasing altitude, and the decreasing rate is different in different height range. The variation of ZTD doesn't vary significantly at different geographic location, and a further study shows that it doesn't have obvious difference at different time too.

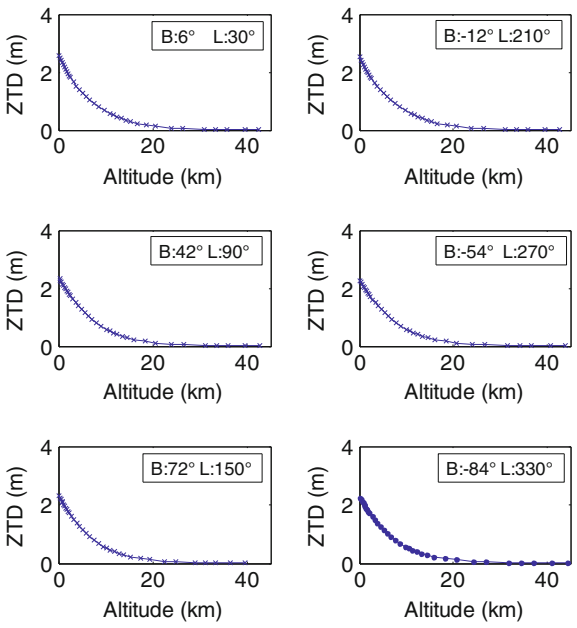


Fig. 36.2 ZTD variation with altitude

36.4.2 Discussion on the Fitting Method of ZTD with Altitude

In the SHAO model, a quadratic polynomial is applied to calculate the adjustment of ZTD in the vertical direction below 10 km, whose accuracy meets the requirement of the surficial navigation users completely. The quadratic polynomial fitting function was used to fit the ZTDs estimated above at all the grid point of ERA-Interim. Figure 36.3 shows the differences between the fitting and integral ZTDs of all the grid point at 12:00 UTC, July 11, 2004. It can be seen that the differences present a regular trend and the worst value of deviation even reach 5 cm. A further statistics for other time also reveals the similar result. It's obvious that there is something unreasonable to apply the quadratic polynomial to fit the variation of ZTD with altitude, so it's necessary for us to find a new approach, which will have more stable precision and wider application range.

On the basis of the diversity of ZTD decreasing rate below 45 km, a piecewise function was introduced to fit the change of ZTD with altitude for the first time. For the range below 5 km, the range between 5 and 10 km, two cubic polynomial functions were applied respectively while an exponential function was applied above 15 km, which can be expressed by the following formulas:

$$ZTD = \sum_{\mathrm{i}=0}^{3} \mathrm{a_{1i}} \times h^i \quad (h < 5\,\mathrm{km}) \tag{36.8}$$

$$ZTD = \sum_{\mathrm{i}=0}^{3} a_{2i} \times h^i \quad (5\,\mathrm{km} < h < 15\,\mathrm{km}) \tag{36.9}$$

$$ZTD = \mathrm{ea} \times \exp(\mathrm{eb} \times \mathrm{h}) \quad (h > 15\,\mathrm{km}) \tag{36.10}$$

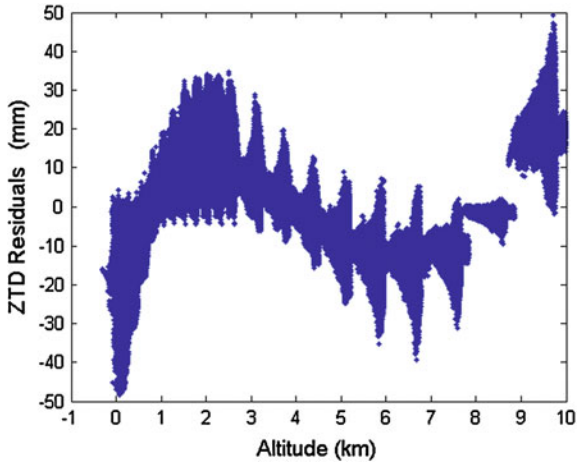


Fig. 36.3 Residuals of fitting the ZTD variation with altitude by quadratic polynomial (below 10 km)

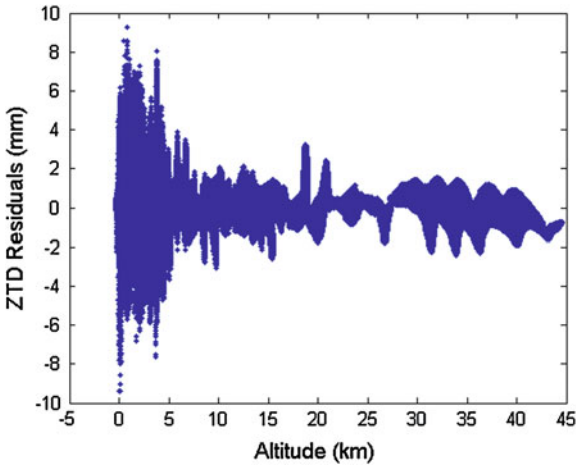


Fig. 36.4 Residuals of fitting the ZTD variation with altitude by piecewise function

The residuals between the piecewise fitting and integral ZTDs of all the grid point at 12:00 UTC, July 11, 2004 are shown in Fig. 36.4. It is can be seen that the residuals almost distribute nearby 0 cm uniformly, and all of them are within 1 cm, especially the deviations above 5 km are less than 0.5 cm. A further study for other time also indicates the similar result, so we can draw a conclusion that the piecewise fitting function mentioned above is a more reasonable choose for the fitting of ZTD and altitude.

36.5 Assessment of Precision

A further comparison between the piecewise fitting ZTDs and the integral ZTDs was implemented so as to evaluate the accuracy of the piecewise fitting function more carefully. In order to assess the spatial distribution of the fitting accuracy, the residuals of the two ZTD sets were group by latitude and longitude band respectively, and the ZTDs come from 45 IGS stations all over the world from 2008 to 2010 were also applied to verify the precision in practical application.

36.5.1 Accuracy Characteristics with Longitude

As shown in Fig. 36.5, the residuals of all the grid points at 12:00 UTC, July 11, 2004 who have the same longitude were gathered together. It can be seen that there is no significant difference between the residual plots of different longitude band, which indicates that the variation of the accuracy with longitude is not obvious. Further statistics for other time also show the same conclusion.

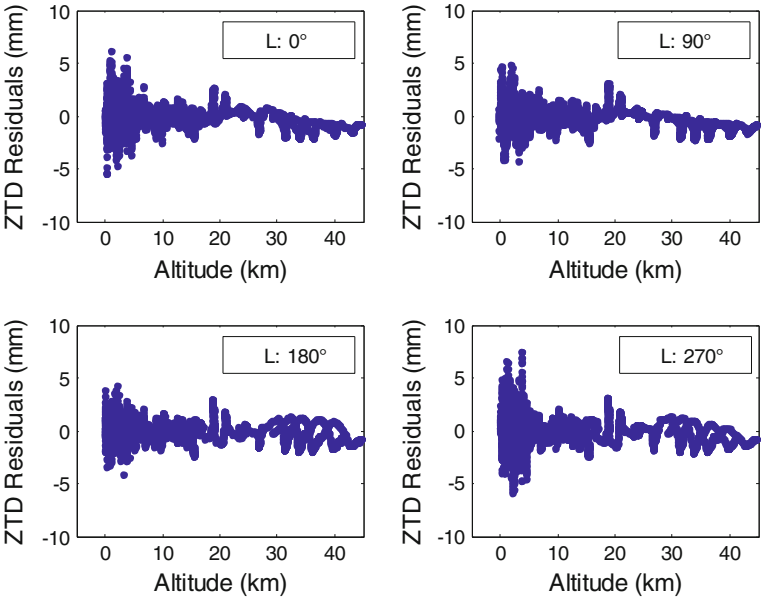


Fig. 36.5 Residuals between the fitting and the integral ZTDs at different longitude

36.5.2 Accuracy Characteristics with Latitude

Figure 36.6 shows the residuals of all the grid points at 12:00 UTC, July 11, 2004 with different latitude. It is obvious that the residuals decrease with the increasing latitude, and the distribution of this trend is symmetrical on the north and south hemispheres. The deviations of the two ZTD sets decrease from 8 mm scale in low-latitude area to 2 mm scale in high-latitude area. Further statistics for other time also show the same result.

36.5.3 Comparison with IGS ZTD

In order to verify the accuracy and applicability of the fitting function in the practical application, after the parameters of the piecewise fitting function were derived, time series of ZTD at the location of 45 IGS stations from 2008 to 2010 were calculated by them. Figure 36.7 shows the two ZTD time series of some stations, and the mean, the maximum and the minimum value of yearly bias and SD were displayed in Table 36.1. The comparison between the fitting and IGS ZTD time series at 45 stations shows a mean yearly difference of no more than 1 cm with a mean SD of about 4.3 cm and the precision of different year is consistent.

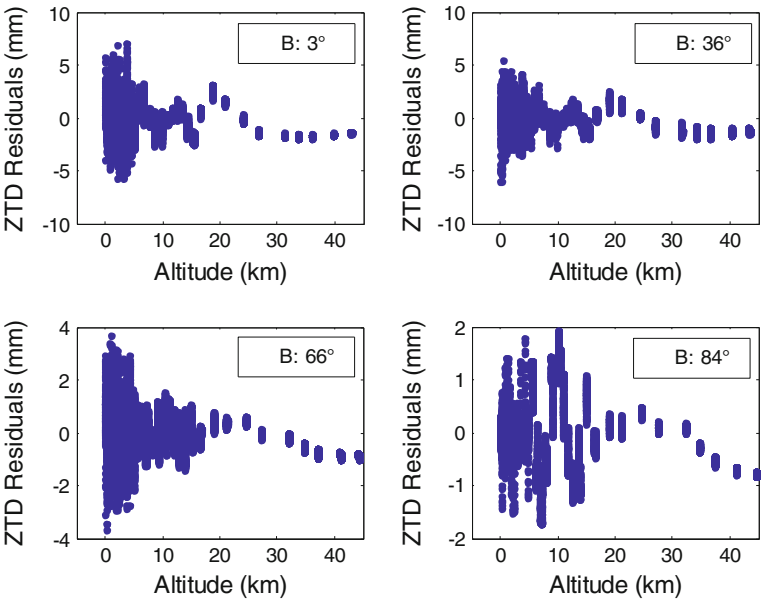


Fig. 36.6 Residuals between the fitting and the integral ZTDs at different latitude

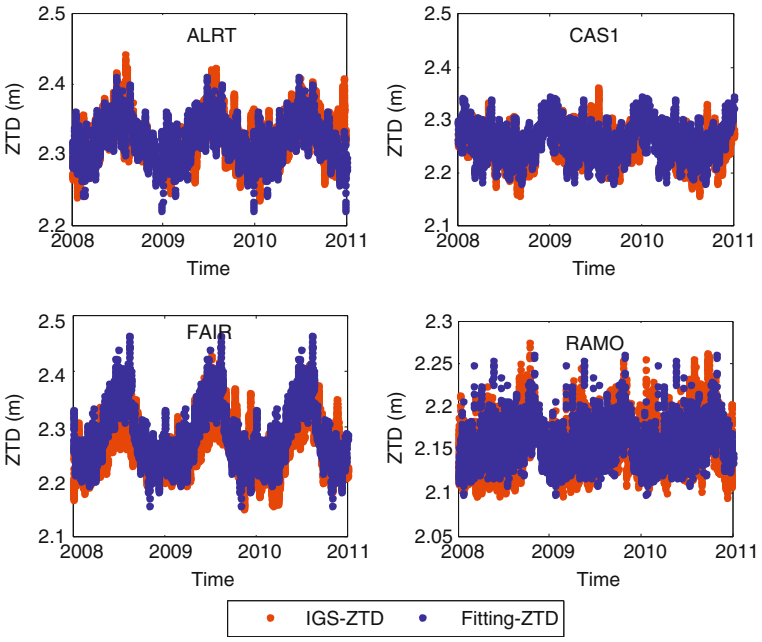


Fig. 36.7 Time series of ZTD

Table 36.1 Year average statistic of bias and SD

	year	Mean	Maximum	Minimum
Bias (cm)	2008	0.51	2.42	0.11
	2009	0.48	3.09	0.07
	2010	0.15	2.61	0.01
SD (cm)	2008	4.39	9.19	2.43
	2009	4.27	8.64	2.41
	2010	4.41	9.16	2.44

36.6 Conclusions

The ZTDs derived from ERA-Interim reanalysis data in 2004 were applied to study the variability of ZTD with height detailedly, and on the basis of that, a new approach of fitting between ZTD and altitude was carried out by a piecewise function. The accuracy of the fitting function above was assessed carefully. The following features of this function can be drawn:

1. It is simple and convenient for the fitting function to be used to estimate ZTD because the function models the ZTD with the altitude directly and the calculation is independent on other external parameters such as meteorological parameters. The ZTDs derived by the function and the IGS ZTDs comparisons at 45 stations from 2008 to 2010 show a mean yearly difference of no more than 1 cm with a mean SD of about 4.3 cm.
2. This function gives a modified method modeling ZTD with attitude to the former models, and the usable range has been extended to 45 km high, which almost includes whole of the troposphere and stratosphere. The parameters are given every 1.5° grid point in the global scale, which will play an important role in the reflection of ZTD in regional area.
3. The accuracy of the function doesn't vary significantly with longitude and time, but has an increasing with the latitude increasing on both the north and south hemispheres.

In conclusion, the piecewise fitting function we discussed above can satisfy the application of real-time navigation and positioning users, and it could be service as a early form of a model, which will have a great value in various navigation area such as the guidance of Low-altitude aircraft.

Acknowledgments Data were provided by ECMWF and IGS. This study was supported by funding from National Nature Science Foundation of China (No. 10603011), National High Technology Research and Development Program 863 (863 Program) (No. 2009AA12Z307).

References

1. Collins JP, Langley RB (1997) A tropospheric delay model for the user of the Wide area augmentation system, final contract report for Nav Canada, Department of Geodesy and Geomatics Engineering Technical Report No 187, University of New Brunswick, Fredericton, NB, Canada
2. Leandro RF, Santos MC, Langley RB (2006) UNB neutral atmosphere models: development and performance. In: ION NTM 2006, Monterey, California, USA, 18–20 Jan 2006, pp 564–573
3. Leandro RF, Langley RB, Santos MC (2008) UNB3m_pack: a neutral atmosphere delay package for radiometric space techniques. GPS Solution 12:65–70
4. Leandro RF, Santos MC, Langley RB (2009) A North America wide area neutral atmosphere model for GNSS applications. Navigation 56:57–71
5. Dodson AH, Chen W, Baker HC et al (1999) Assessment of EGNOS tropospheric correction model. In: ION GPS 1999, Nashville, Tennessee, USA, 14–17 Sept 1999, pp 1401–1407
6. Penna N, Dodson A, Chen W (2001) Assessment of EGNOS tropospheric correction model. J Navig 54:37–55
7. Uemo M, Hoshinoo K, Matsunaga K et al (2001) Assessment of atmospheric delay correction models for the Japanese MSAS. In: ION GPS 2001, Salt Lake City, Utah, USA, 11–14 Sept 2001, pp 2341–2350
8. Song SL, Zhu WY, Chen QM et al (2011) Establishment of a new tropospheric delay correction model over China area. Sci China 54(12):2271–2283
9. Chen QM, Song SL, Zhu WY (2011) Preliminary establishment of the global model (SHAO-G) for the zenith tropospheric delay. In: China Satellite Navigation Conference, Shanghai
10. Dee DP, Uppala SM, Simmons AJ et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol 137:553–597
11. Chen QM, Song SL, Heise S et al (2011) Assessment of ZTD derived from ECMWF/NCEP data with GPS ZTD over China. GPS Solution 15(4):415–425
12. Chen QM, Song SL, Zhu WY (2011) An analysis of the accuracy of zenith tropospheric delay calculated from ECMWF/NCEP data over Asian area. Chin J Geophys 55(5):1541–1548

Chapter 37
COMPASS Three Carrier Ambiguity Resolution

Zhiqiang Dai, Qile Zhao, Zhigang Hu, Xing Su, Lizhong Qu and Qiang Guo

Abstract The new generation of global navigation satellite systems (GNSSs) introduce a third (or more) carrier, which benefit ambiguity resolution (AR). TCAR (Three-Carrier Ambiguity Resolution) and CIR (Cascaded Integer Resolution) are the typical three-carrier ambiguity resolution methods, which are biased by the residual ionospheric delay. Based on the COMPASS three-carrier observables, this contribution investigated the TCAR/CIR methods, and accommodated them to the middle and long baseline by introducing some modifications. The modified TCAR/CIR methods contributes to the second and third steps of the old methods, which reduces the error in the second step slightly from 0.547 to 0.478, and that in the third step sharply form 17.913 to 5.062.

Keywords GNSS · TCAR · CIR · COMPASS/BeiDou-2

37.1 Introduction

The new generation of GNSSs all transmit three or more carriers, such as the modernized GPS introduce the L5 signal, besides the previous L1 and L2 signals, the ongoing European Galileo system is designed to transmit L1, E6, E5B, E5A four signals, and the in-operation Chinese COMPASS/Beidou-2 navigation system is transmitting E2, E6, E5B three signals.

Z. Dai (✉) · Z. Hu
School of Geodesy and Geomatics, Wuhan University, No. 129, Luoyu Road, Wuchang District, Wuhan, Hubei, People's Republic of China
e-mail: dzq@whu.edu.cn

Q. Zhao · X. Su · L. Qu · Q. Guo
National Engineering Research Center for Satellite Navigation System, Wuhan University, No. 129, Luoyu Road, Wuchang District, Wuhan, Hubei, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_37,

Multi-frequency carriers not only introduce more observables, but also form more linear combinations between the observables, which contribute to the AR. The TCAR [1, 2] and CIR [3, 4] are the typical three-carrier ambiguity resolution methods, which are equivalent and both the geometry-free bootstrapping method [5]. TCAR and CIR have the same principle, i.e. first initial with the extra wide-lane (EWL) ambiguity resolution, then resolve the wide-lane (WL) ambiguity based on the unambiguous EWL observable, at last resolve the narrow-lane (NL) ambiguity using the unambiguous EWL and WL observables. Although TCAR/CIR can resolve the EWL ambiguity with high accuracy and efficience, the resolution of WL and NL ambiguity would be biased by the residual ionospheric delay for the shorter wavelengths of WL and NL observables. To dealt with the residual ionospheric delay, Feng Y. proposed a geometry-based TCAR method using the ionosphere-reduced virtual signals [6]. This method firstly generates the virtual observables from linear combinations of the original observables which is biased by a minimum ionospheric error, and then resolve the ambiguity using geometry-based model and searching algorithm. Although the residual ionospheric delay is reduced, it becomes larger when the baseline length increases and the spacial correlation decreases. Thus this method is suit for the short baseline only. Meanwhile, the searching algorithm could reduce the efficience.

The previous three-carrier ambiguity resolution researches focus on the theory and simulation, for there are not enough GNSS satellites transmitting three (or more) signals. The Chinese COMPASS/BeiDou-2 system consists of 5 GEO satellites, 5 IGSO satellites and 4 MEO satellites, and there are more than 10 satellites can transmit B1, B3 and B2 signals, which can be used to form a extra wide-lane and two wide-lane observables. Table 37.1 lists the original and (extra) wide-lane signals of COMPASS/BeiDou-2 satellites. The work in this contribution is based on the COMPASS three-frequency observables.

37.2 Three-Carrier Ambiguity Resolution

TCAR and CIR are the typical methods for three-carrier ambiguity resolution. We first introduced and analyzed these methods, and then introduced some modifications to accommodate them to the middle and long baseline. In the following introduction, the carrier and code observables are double-differenced in

Table 37.1 Original and (extra) wide-lane signals of COMPASS satellites

Signal	Frequency (MHz)	Wavelength (m)
B1	1561.098	0.192
B3	1268.520	0.236
B2	1207.140	0.248
B1–B2	353.958	0.847
B1–B3	292.578	1.025
B3–B2	61.380	4.884

default (omitting the $\Delta\nabla$ operator for simplicity), and in meter. The tokens f_1, f_2 and f_3 ($f_1 > f_2 > f_3$) represent the frequencies of B1, B3 and B2 signals, L_1, L_2 and L_3 represent the carrier phase observables, and P_1, P_2 and P_3 represent the code observables, respectively. In the analysis, there is an assumption that the fixed ambiguity in each step is unbiased, which is used as a true value in the following step.

37.2.1 TCAR/CIR Methods

TCAR/CIR methods firstly generate the extra wide-lane and wide-line observables, which are sorted by the wavelength, then initial with the M-W combination [7, 8] to determine the EWL ambiguity, and then step to the shorter WL ambiguity resolution, and at last the NL ambiguities are resolved. In this procedure, the later step must use the result of the former step.

The double-differenced phase and code observables can be represented as

$$\begin{cases} L_i = \rho' - \eta_i I_1 - \lambda_i N_i + \varepsilon_L \\ P_i = \rho' + \eta_i I_1 + \varepsilon_P \end{cases}, \ \eta_i = \frac{f_1^2}{f_i^2}, \ i = 1, 2, 3$$

where L_i and P_i are the double-differenced phase and code observables, ρ' contains the double-differenced geometric distance and non-dispersive errors, I_1 is the double-differenced ionospheric delay in L_1 and P_1, η_i is the scale factor of residual ionospheric delay in L_1 and P_i with regard to L_1 and P_1, λ_i and N_i are the wavelength and ambiguity of L_i, ε_L and ε_P are noise in carrier and code observables with the assumption $\sigma_{L_1} = \sigma_{L_2} = \sigma_{L_3} = \sigma_L$ and $\sigma_{P_1} = \sigma_{P_2} = \sigma_{P_3} = \sigma_P$. For the COMPASS/BeiDou-2 system $\sigma_L = 0.005\,\text{m}$, $\sigma_P = 0.8\,\text{m}$ [9].

There is combination between any two observables:

$$\begin{cases} L_{ab} = \dfrac{f_a L_a - f_b L_b}{f_a - f_b} = \rho' + \eta_{ab} I_1 - \lambda_{ab} N_{ab} + \varepsilon_{L_{ab}} \\ P_{ab} = \dfrac{f_a P_a + f_b P_b}{f_a + f_b} = \rho' + \eta_{ab} I_1 + \varepsilon_{P_{ab}} \end{cases}, \ \eta_{ab} = \frac{f_1^2}{f_a f_b}, \ a, b = 1, 2, 3 \ and \ a \neq b$$

where λ_{ab}, N_{ab} and η_{ab} denote the wavelength, ambiguity and residual ionospheric scalar factor of EWL and WL observables L_{ab}. Then the TCAR/CIR express as the following steps:

1. Resolve the EWL ambiguity $\hat{N}_{23}$

$$N_{23} = (P_{23} - L_{23} + \varepsilon_{23})/\lambda_{23}$$
$$\sigma_{23} = \sigma_P \sqrt{f_2^2 + f_3^2} \Big/ (f_2 + f_3)$$
$$\hat{N}_{23} = Int(N_{23})$$

2. Resolve the WL ambiguity $\hat{N}_{12}$ and $\hat{N}_{13}$

$$N_{12} = \left[L_{23} + \lambda_{23}\hat{N}_{23} - L_{12} - (\eta_{23} - \eta_{12})I_1 + \varepsilon_{12}\right]/\lambda_{12}$$
$$\eta_{23} - \eta_{12} = \frac{f_1(f_3 - f_1)}{f_2 f_3}, \sigma_{12} = \sqrt{\sigma_{L_{23}}^2 + \sigma_{L_{12}}^2}$$
$$\hat{N}_{12} = Int(N_{12}),\ \hat{N}_{13} = \hat{N}_{12} + \hat{N}_{23}$$

3. Resolve the NL ambiguity $\hat{N}_1$, $\hat{N}_2$ and $\hat{N}_3$

$$N_3 = \left[L_{13} + \lambda_{13}\hat{N}_{13} - L_3 - (\eta_{13} + \eta_3)I_1 + \varepsilon_3\right]/\lambda_3$$
$$\eta_{13} + \eta_3 = -\frac{f_1(f_1 + f_3)}{f_3^2},\ \sigma_3 = \sqrt{\sigma_{L_{13}}^2 + \sigma_{L_3}^2}$$
$$\hat{N}_3 = Int(N_3),\ \hat{N}_2 = \hat{N}_{23} + \hat{N}_3,\ \hat{N}_1 = \hat{N}_{13} + \hat{N}_3$$

where N_{ab} and N_i are the float ambiguities, ε_{ab} and ε_i denote the noises of ambiguities N_{ab} and N_i, whose variances are σ_{ab} and σ_i, $\hat{N}_{ab}$ and $\hat{N}_i$ are the fixed ambiguities.

Of the three steps, the EWL ambiguity $\hat{N}_{23}$ in step (1) is affected by the measurement noise, while the WL ambiguity $\hat{N}_{12}$ in step (2) and NL ambiguity $\hat{N}_3$ in step (3) are biased by the residual ionospheric error in addition to the measurement noise. It is obvious that TCAR/CIR methods ignore the residual ionospheric error, which limits their application to short baseline. As the baseline length increases, the spacial correlation of ionospheric error decreases and there is larger residual ionospheric delay. Figure 37.1 shows the residual ionospheric delay in Changsha-Guangzhou baseline (563 km) for different satellite pairs. The blue line represent the residual ionospheric delay derived from the unambiguous WL observables L_{12} and L_{13}, and the red line represent the residual ionospheric delay derived from the unambiguous original observables L_1 and L_2. As it is shown, the

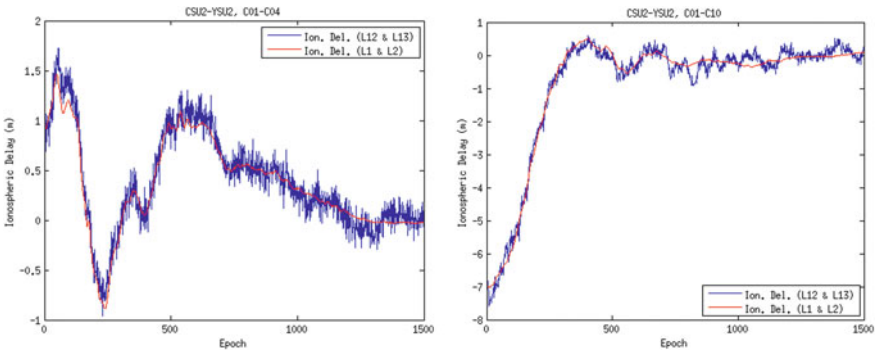


Fig. 37.1 The residual ionospheric delay in Changsha-Guangzhou 563 km baseline for satellite pairs C01-C04 and C01-C10

Table 37.2 Errors in the EWL, WL and NL ambiguities resolution using TCAR/CIR method

Ambiguity	Error expression	Error quantity (cycle)
$\hat{N}_{23}$	$\sigma_P \sqrt{f_2^2 + f_3^2} \Big/ \lambda_{23}(f_2 + f_3)$	0.116
$\hat{N}_{12}$	$\sqrt{\sigma_{12}^2 + (\eta_{23} - \eta_{12})^2 \sigma_{I_1}^2} \Big/ \lambda_{12}$	0.547
$\hat{N}_3$	$\sqrt{\sigma_3^2 + (\eta_{13} + \eta_3)^2 \sigma_{I_1}^2} \Big/ \lambda_3$	17.913

residual ionospheric delay is a few meters and varies sharply, which makes the ambiguity resolution more difficult.

Table 37.2 lists the errors in the above procedures. We assume the variance of COMPASS code and phase observables are 0.8 and 0.005 m respectively [9], and the residual ionospheric error in L_1 and P_1 is 1.5 m. As it is showed in Table 37.2, there is large error in the TCAR/CIR method, i.e. more than 0.5 cycles in step (2) and about 18 cycles in step (3), which hamper the ambiguity resolution.

37.2.2 Modified TCAR/CIR Method

To deal with the residual ionospheric delay in middle and long baseline, we modified the step (2) and (3) of TCAR/CIR, eliminating the residual ionospheric delay. The key point of the modification is forming the linear combination between code observables or unambiguous EWL and WL observables. The linear combination coefficients should be subject to the conditions: minimal noise, unchanged geometry scale and identical ionospheric delay with that of the observable to be resolved the ambiguity.

Using P_1, P_3 and the unambiguous L_{23} in step (1), we can get the combined observable as follow:

$$\widehat{L}_{12} = \alpha \cdot P_1 + \beta \cdot P_2 + \gamma \cdot (L_{23} + \lambda_{23}\hat{N}_{23})$$

where the coefficients α, β and γ subject to:

$$\begin{cases} \alpha + \beta + \gamma = 1 \\ \alpha\eta_1 + \beta\eta_2 + \gamma\eta_{23} = \eta_{12} \\ (\alpha^2 + \beta^2)\sigma_P^2 + \gamma^2\sigma_{L_{23}}^2 = \min \end{cases}$$

Thus we get $\alpha = 0.599$, $\beta = -0.082$ and $\gamma = 0.483$ for the current COMPASS signals. Then the step (2) can be modified as:

Table 37.3 Errors in the EWL, WL and NL ambiguities resolution using modified TCAR/CIR method

Ambiguity	Error expression	Error quantity (cycle)
$\hat{N}_{23}$	$\sigma_P\sqrt{f_2^2+f_3^2}\Big/\lambda_{23}(f_2+f_3)$	0.116
$\hat{N}_{12}$	$\sqrt{\sigma^2_{\widehat{L}_{12}}+\sigma^2_{L_{12}}}\Big/\lambda_{12}$	0.478
$\hat{N}_3$	$\sqrt{\sigma^2_{\widehat{L}_3}+\sigma^2_{L_3}}\Big/\lambda_3$	5.062

$$
\begin{aligned}
N_{12} &= \left[\widehat{L}_{12} - L_{12} + \varepsilon_{12}\right]\Big/\lambda_{12} \\
\sigma_{12} &= \sqrt{\sigma^2_{\widehat{L}_{12}} + \sigma^2_{L_{12}}} \\
\hat{N}_{12} &= Int(N_{12}),\ \hat{N}_{13} = \hat{N}_{12} + \hat{N}_{23}
\end{aligned}
$$

Similarly, using the unambiguous L_{23} in step (1), L_{12} and L_{13} above, we can get the combined observable as follow:

$$\widehat{L}_3 = \alpha\cdot(L_{23}+\lambda_{23}\hat{N}_{23}) + \beta\cdot(L_{12}+\lambda_{12}\hat{N}_{12}) + \gamma\cdot(L_{13}+\lambda_{13}\hat{N}_{13})$$

where α, β and γ have the similar restriction:

$$
\begin{cases}
\alpha+\beta+\gamma = 1 \\
\alpha\eta_{23} + \beta\eta_{12} + \gamma\eta_{13} = -\eta_3 \\
\alpha^2\sigma^2_{L_{23}} + \beta^2\sigma^2_{L_{12}} + \gamma^2\sigma^2_{L_{13}} = \min
\end{cases}
$$

Thus, we get $\alpha = -8.661$, $\beta = 6.111$ and $\gamma = 3.550$. The step (3) can be modified as:

$$
\begin{aligned}
N_3 &= \left[\widehat{L}_3 - L_3 + \varepsilon_3\right]\Big/\lambda_3 \\
\sigma_3 &= \sqrt{\sigma^2_{\widehat{L}_3} + \sigma^2_{L_3}} \\
\hat{N}_3 &= Int(N_3),\ \hat{N}_2 = \hat{N}_{23} + \hat{N}_3,\ \hat{N}_1 = \hat{N}_{13} + \hat{N}_3
\end{aligned}
$$

The above modifications eliminate the residual ionospheric delay, geometry distance and the non-dispersive error. As a result, there is only the random noise, which can be reduced by averaging multiple observables, though it is magnified by the modification.

Table 37.3 list the errors in AR procedure using modified TCAR/CIR method. We can find the modified TCAR/CIR method reduces the errors in step (2) slightly from 0.547 to 0.478, and that in step (3) sharply form 17.913 to 5.062.

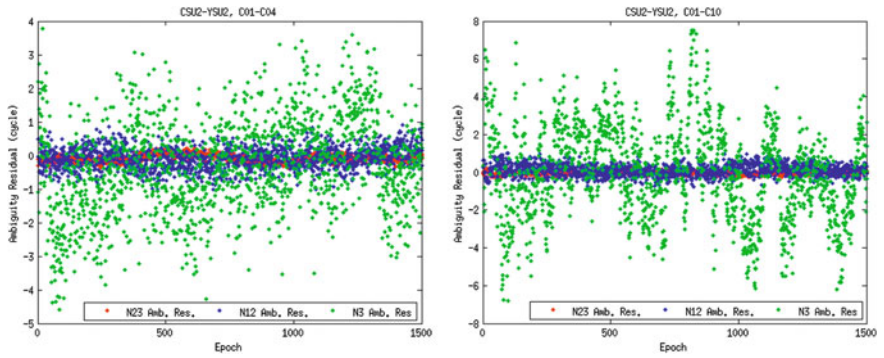


Fig. 37.2 EWL, WL and NL ambiguity residuals in Changsha-Guangzhou 563 km baseline for different satellite pairs C01-C04 and C01-C10

37.3 Result

Based on the synchronous COMPASS observables in Changsha-Guangzhou (563 km) baseline from 8:45 (UTC), Nov 23, 2012 to 21:15 (UTC), Nov 23, 2012, we validated the modified TCAR/CIR method.

Adopting the modified TCAR/CIR method proposed by this contribution, we resolved the EWL, WL and NL ambiguities. The ambiguity residuals (the difference between each float ambiguities and the fixed ambiguity) are display in Fig. 37.2. The ambiguity residuals of EWL are less than 0.5 cycle, which means the EWL ambiguity can be resolved at a single epoch. The WL ambiguity residuals are less than 1 cycle, so the WL ambiguity can be resolved by averaging a few epochs. As the NL ambiguity residuals exceed to a few cycles, the NL ambiguity should be resolved by averaging much more observables in a long period.

37.4 Conclusion and Discussion

The TCAR and CIR methods both ignore the double-differenced residual ionospheric delay, which restricts their application to short baseline (less than 10 km). For the middle and long baseline, residual ionospheric delay could introduce large systematic error, which biases the ambiguity resolution. To deal with this problem, this contribution modified the TCAR/CIR method, resulting in a ionosphere- and geometry-free three-carrier ambiguity resolution method. The modified TCAR/CIR method reduces the errors in step (2) slightly and that in step (3) sharply. For the 563 km long baseline, the proposed method can resolve the EWL ambiguity at a single epoch, and the WL ambiguity by averaging a few observables, while the NL should be resolved by averaging much more observables in a long period.

References

1. Forssell B, Martin-Neira M, Harris AR (1997) Carrier phase ambiguity resolution in GNSS-2. In: Proceeding of ION GPS-97, Kansas City, MO, 16–19 Sept 1997
2. Vollath U, Birnbach S, Landau H (1998) Analysis of three carrier ambiguity resolution (TCAR) technique for precise relative positioning in GNSS-2. In: Proceeding of ION GPS 1998, pp 417–426
3. De Jonge PJ, Teunissen PJG, Jonkman NF, Joosten P (2000) The distributional dependence of the range on triple frequency GPS ambiguity resolution. In: Proceedings of ION-NTM 2000, Anaheim, CA, 26–28 Jan 2000, pp 605–612
4. Hatch R, Jung J, Enge P, Pervan B (2000) Civilian GPS: the benefits of three frequencies. GPS Solution 3(4):1–9
5. Teunissen PJG, Joosten P, Tiberius C (2002) A comparison of TCAR, CIR and LAMBDA GNSS ambiguity resolution. In: Proceedings of ION GPS, Portland, Oregon, 24–27 Sept 2002, pp 2799–2808
6. Feng Y (2008) GNSS three carrier ambiguity resolution using ionosphere-reduced virtual signals. J Geodesy 82:847–862
7. Melbourne WG (1985) The case in GPS based geodetic systems. First symposium on precise positioning with global positioning system, Rockville, Maryland, USA, pp 373–386
8. Wübbena G (1989) The GPS adjustment software package GEONAP: concepts and models. In: Proceedings of 5th International Geodesy Syrup on satellite positioning. Las Cruces, New Mexico, 13–17 Mar 1989, pp 452–461
9. Oliver M et al (2012) Initial assessment of the COMPASS/BeiDou-2 regional navigation satellite system, GPS Solution 2012. doi:10.1007/s10291-012-0272-x

Chapter 38
GNSS Precise Point Positioning Algorithm Based on Parameter Equivalent Reduction Principle

Guanwen Huang, Yuanxi Yang, Chao Liu, Qin Zhang and Shuangcheng Zhang

Abstract The standard GNSS combined PPP algorithm cannot be used satisfactorily in the real-time and high frequency precise positioning because of its low compute efficiency. A new algorithm based on the parameter equivalent reduction principle is proposed. First, the observations equation and the normal equation which belong to the single navigation system can be solved independently. Second, the normal equations of overlapping parameters between the different systems can be obtained by using parameter equivalent reduction principle. At last, the combined PPP resolutions can be computed easily by using the Least Squares method. The proposed algorithm can improve the calculating efficiency immensely. In addition, an adaptively combined method which can automatically adjusts the contributed weight of different GNSS systems is also proposed in this paper. The numerical examples using the data set of three IGS stations, show that the PPP precisions and efficiencies based on the proposed model have been improved significantly compared with those of standard model. This proposed principle can also be applied in the GNSS precise satellite clock determination and the indifference baseline network adjustment plus the GNSS time offset monitoring.

Keywords Parameter reduction · PPP · Combined positioning · GPS/GLONASS

G. Huang (✉) · Q. Zhang · S. Zhang
College of Geology Engineering and Geomantic, Chang'an University,
Xi'an 710054 Shaanxi, China
e-mail: huang830928@yahoo.com.cn

Y. Yang
China National Administration of GNSS and Applications, Beijing 100088, China

Y. Yang
State Key Laboratory of Geographic Information Engineering, Yanta Road 126,
Xi'an 710054, China

C. Liu
College of Geology Engineering and Geomantic, Chang′an University,
Xi′an 710054 Shaanxi, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_38,

38.1 Introduction

With the official start using of Chinese BDS satellite navigation system, the amount of global satellite navigation system (GNSS) which are providing services has increased from two (GPS of America and GLONASS of Russia) to three (GPS/GLONASS/BDS). There are more than 70 worked satellites now and with the complete of European Galileo system (Galileo), the total amount of GNSS satellites will exceed 100, the majority of which are multi-frequency worked satellites. GNSS has proceeded to a new age of new style with multiple constellations and frequencies [1, 2].

With the modernized development of GNSS, the precision of multiple systems combined navigation positioning and timing is higher and higher. The developing characters of GNSS have been changed gradually from the single GPS positioning with posterior relative positioning in the early period to the multiple systems GNSS combined absolute positioning with the real-time and high-frequency observations. Undoubtedly, the rapid and flexible Precise Point Positioning (PPP) will be one of the most prospective application technologies [3–9].

The precision and stability of real-time rapid positioning will be improved immensely by using GNSS multiple systems combined positioning [1, 2]. In recent years, there are abundant researches results about GNSS combined PPP technologies at domestic and abroad [10–17], which mainly focus on GPS/GLONASS combination. However, these results are mainly demonstrating the precision and reliability advantages brought by combined positioning and the combined algorithm itself and the contributed weight distribution of different systems are involved rarely.

The traditional combined PPP algorithm directly constitutes observation equation using all GNSS observations to obtain the solution. With the increasing of available GNSS systems and satellites, and the widely applying of real-time high-frequency (1–50 Hz) receiver data [3], the processing burdens of the traditional algorithm will increase exponentially with the addition of GNSS systems. In addition, the traditional method cannot adaptively adjust the contributed weights of the different systems. It will influence the precision and reliability of the combined PPP positioning results. Some improved methods are proposed for the satellite clock estimating and GNSS network adjustment which use the Parameter Equivalent Reduction Principle to reduce the number of the unknown parameters and enhance the computational efficient [6]. Parameter Equivalent Reduction Principle also is used in the several famous GNSS software such as GAMIT/GLOBK and Bernese.

In there, a new GNSS combined PPP algorithm which using sequential PPP algorithm based on equivalently eliminated parameter principle [18–21] is proposed. First the whole combined observations of multi-models PPP is divided into each single system to be solved independently and paralleled. Second the normal equation matrixes of overlapped parameters which belong to the different systems have been accumulated. Third the combined resolution can be obtained directly by

solving the accumulated normal equation matrix. The new algorithm improves the calculating efficiency of the GNSS combined Positioning immensely. In the meanwhile, a combined positioning method which uses posterior variance factors to adaptively adjust the contributed weight relations of different GNSS systems is also proposed. At last, the availability and precision of the new algorithm are verified by analysis of different algorithms' calculating time and the PPP positioning precision.

38.2 The Traditional GNSS Combined PPP Algorithm

38.2.1 The Standard PPP Observation Equation

The conventional un-different Precise Point Positioning method usually uses ionosphere-free combined observations of carrier phases and pseudoranges. The observation equations as follow:

$$\Phi_{IF} = \frac{f_1^2\Phi_1 - f_2^2\Phi_2}{f_1^2 - f_2^2} = \rho + c(dt_r - dt_s) + d_{trop} + \lambda N + \delta m_{IF} + \varepsilon(\Phi_{IF}) \quad (38.1)$$

$$P_{IF} = \frac{f_1^2 \cdot P_1 - f_2^2 \cdot P_2}{f_1^2 - f_2^2} = \rho + c(dt_r - dt_s) + d_{trop} + dm_{IF} + \varepsilon(P_{IF}) \quad (38.2)$$

where, P_{IF} and Φ_{IF} are the ionosphere-free combinations of carrier phases and pseudoranges respectively. f_1 and f_2 are the different frequency of dual-frequency observations; ρ is the geometric distance between station and satellite; dt_r denotes the clock error of receiver; dt_s denotes the clock error of satellite; dt_{trop} denotes the tropospheric delay; δm_{IF} and dm_{IF} denote a sequence of error corrections including relativistic effects, earth tide, solid tide and hardware delay. λ is the wavelength of combined carrier; N is the combined ambiguity. $\varepsilon(P_{IF})$ and are the noise errors of code combination and phase combination observations respectively.

The observation equation of (38.1) and (38.2) is as follow:

$$V = AX - L,\ P \quad (38.3)$$

where A is the observation coefficient matrix; X is the estimated parameter matrix, including coordinates parameter, troposphere parameter and ambiguity parameter; L is the observation matrix; V is the observation's residual matrix; P is the observation's weight matrix.

The sequential least squares adjustment and kinematic Kalman filter [9] are used to solve (38.3) generally. The sequential least squares adjustment is suitable for static positioning while the kinematic Kalman filter is suitable for kinematic positioning. For example, the sequential least squares PPP solutions can be described using Eq. (38.3) as follow:

$$\begin{cases} X_i = \left(A_i^T P_i A_i + Q_{i-1}^{-1}\right)^{-1}\left(A_i^T P_i L_i + Q_{i-1}^{-1} X_{i-1}\right) \\ Q_i = \left(A_i^T P_i A_i + Q_{i-1}^{-1}\right)^{-1} \end{cases} \tag{38.4}$$

where Q_{i-1} is the covariance matrix of unknown parameters of epoch i − 1; X_{i-1} is unknown parameters vectors of epoch i − 1; Q_i and X_i are the covariance matrix of unknown parameters of epoch i and unknown parameters vectors of epoch i.

The above equations are the PPP observation equation and its sequential solution formula of single navigation positioning system. The observation equation and resolution formula of the conventional GNSS multi-systems in combination with PPP positioning will be discussed below.

38.2.2 The Traditional Combined PPP Algorithm

At present, there are three GNSS systems which have been widely used, including GPS, GLONASS and BDS system. Taking these three systems for example, the observation equation of GNSS in combination with PPP is as follow:

$$\begin{bmatrix} V^G \\ V^R \\ V^C \end{bmatrix} = \begin{bmatrix} A^G \\ A^R \\ A^C \end{bmatrix} X - \begin{bmatrix} L^G \\ L^R \\ L^C \end{bmatrix}, \begin{bmatrix} P^G & 0 & 0 \\ 0 & P^R & 0 \\ 0 & 0 & P^C \end{bmatrix} \tag{38.5}$$

where, the superscripts G, R and C represent GPS, GLONASS and BDS respectively. The X parameters contain coordinate parameters, troposphere parameter, the each receiver's clock parameter of three systems and ambiguity parameters. The geometric meanings of other matrixes are the same to Eq. (38.3). For the solution of Eq. (38.5), the conventional method is to use the sequential formula as Eq. (38.4).

The multi-systems whole arranged observation equation form as Eq. (38.5) has a clear meaning and can be understand easily. However, with the addition of navigation systems, the exponent increasing of calculating burden is harmful to the rapidly processing of large amount of real-time and high-frequency data.

For example, if the multi-models receiver can observe n satellites from m navigation systems at i epoch, the dimension of A matrix will be 2mn × (6+ mn); the dimension of X will be (6+ mn) × 1; the dimension of P will be 2 × 2mn; the dimension of V will be 2mn×1. In the recent GNSS systems, GPS, GLONASS and BDS are widely used. Therefore m can be set as 3. With the good observable sight, the available satellites of each system can achieve 10, then the dimension of A, X, P, V and L will be 60 × 36, 36 × 1, 60 × 60, 60 × 1, 60 × 1 respectively. When the original observations instead of LC observations are used to calculate, the corresponding dimension will expand to about twice. There are hundreds of even thousands of stations in current regional or global GNSS real-time monitoring network, how to improve the efficiency of combined algorithm is one of the important aspects to realize rapid positioning resolution.

To the solution and calculation of above observation matrix, taking the process procedures of preprocessing and iteration calculation in the practical positioning

into consideration, the conventional algorithm that obtains solution by direct accumulation cannot meet the demand of rapid positioning of large amount of high-frequency and real-time data. With the increasing of available satellites, it will bring exponent increasing calculating burden to the conventional positioning algorithm.

On the other hand, the conventional multi-systems PPP model cannot balance the contribution weight of different systems effectively. The precision of whole combined solution will be influenced even the combined solution is mistaken when there are blunders in the observation of one system.

Therefore, the equivalently eliminated parameter principle is proposed to solve the multi-systems combined PPP and to adaptively adjust the combined scale factors of different systems using posterior variance. Then the calculating efficiency of the combined PPP positioning is enhanced. The precision and robustness of the combined solution are improved immensely.

At first, the parameter vector and its covariance estimation formula has been rewritten to integrated form of normal equation coefficient matrix to deliver in each epoch. The specific form is as follow:

$$\begin{cases} N_i = A_i^T P_i A_i + N_{i-1} \\ W_i = A_i^T P_i L_i + W_{i-1} \\ X_i = N_i^{-1} W_i \\ Q_i = N_i^{-1} \end{cases} \tag{38.6}$$

In the above equation, the integrated form of traditional parameter vector X_i and covariance matrix Q_i in each epoch has been transferred into the normal equation coefficient matrix N_i and W_i integration. Compared to the whole solution, the calculating burden has not be increased and the dimension of N_i and W_i is the same to Q_i and X_i. Both are equivalent when processing in real time. In the unreal-time processing (such as post processing), the inversion of Q_i and X_i at each epoch can be omitted and only solve the accumulated normal equation coefficient matrix at last epoch in the normal equation integration. Its calculating efficiency is superior to the traditional parameter and variance matrix integrated method. The model of multi-systems combined PPP algorithm is derived below based on the method of normal equation coefficient matrix inheritance (Eq. 38.6).

38.3 The Multi-systems Combined PPP Algorithm Based on the Parameter Elimination

38.3.1 The Single-System PPP Algorithm Based on the Parameter Elimination

To divide the parameter X in single system observation equation into two kinds, the corresponding divided observation equation can be written as a block matrix form:

$$V = A_1X_1 + A_2X_2 - L,\ P \tag{38.7}$$

where X_1 is the parameter that changes by time, which can be called the changeable parameter. In PPP it mainly indicates the receiver clock error parameter; X_2 is the parameter that won't change by time, which is fixed parameter and it means can be inherited during epochs. In PPP it mainly indicates the coordinate parameter, troposphere parameter and ambiguity parameter. It should be pointed that the coordinate parameters of kinematic PPP positioning during epochs are changeable and can not be inherited directly, which can be predicted through a filter state equation form.

The block equation obtained from Eq. (38.7) is as follow:

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} W_{11} \\ W_{22} \end{bmatrix} \tag{38.8}$$

where, $N_{11} = A_1^TPA_1,\ N_{12} = A_1^TPA_2,\ N_{21} = N_{12}^T,\ N_{22} = A_2^TPA_2,\ W_{11} = A_1^TPL,\ W_{22} = A_2^TPL.$

When Eq. (38.8) is parameter equivalently eliminated [18], the formula below can be obtained.

$$\begin{bmatrix} N_{11} & N_{12} \\ 0 & N_2 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} W_{11} \\ R_2 \end{bmatrix} \tag{38.9}$$

where, $N_2 = N_{22} - N_{21}N_{11}^{-1}N_{12},\ R_2 = W_{22} - N_{21}N_{11}^{-1}W_{11}.$

For Eq. (38.9), X_2 can be obtained from solving the second expression, X_1 can be regained from the first expression as follow:

$$N_2X_2 = R_2 \tag{38.10}$$

$$N_{11}X_1 = W_{11} - N_{12}X_2 \tag{38.11}$$

According to the inherited form of normal equation of Eq. (38.6), the single system PPP sequential formula of i epoch can be obtained as follow:

$$\begin{cases} \hat{\mathrm{N}}_2^{\mathrm{i}} = \mathrm{N}_2^{\mathrm{i}} + \hat{\mathrm{N}}_2^{\mathrm{i}-1} \\ \hat{\mathrm{R}}_2^{\mathrm{i}} = \mathrm{R}_2^{\mathrm{i}} + \hat{\mathrm{R}}_2^{\mathrm{i}-1} \\ \mathrm{X}_2 = \left(\hat{\mathrm{N}}_2^{\mathrm{i}}\right)^{-1}\hat{\mathrm{R}}_2^{\mathrm{i}} \\ \mathrm{X}_1^{\mathrm{i}} = \left(\mathrm{N}_{11}^{\mathrm{i}}\right)^{-1}\left(\mathrm{W}_{11}^{\mathrm{i}} - \mathrm{N}_{12}^{\mathrm{i}}\mathrm{X}_2\right) \\ \mathrm{Q}_{\mathrm{X}_2} = \left(\hat{\mathrm{N}}_2^{\mathrm{i}}\right)^{-1} \\ \mathrm{Q}_{\mathrm{X}_1^{\mathrm{i}}} = \left(\mathrm{N}_{11}^{\mathrm{i}} - \mathrm{N}_{12}^{\mathrm{i}}\left(\mathrm{N}_{22(\mathrm{new})}^{\mathrm{i}}\right)^{-1}\mathrm{N}_{21}^{\mathrm{i}}\right)^{-1}, \mathrm{N}_{22(\mathrm{new})}^{\mathrm{i}} = \hat{\mathrm{N}}_2^{\mathrm{i}} + \mathrm{N}_{21}^{\mathrm{i}}\left(\mathrm{N}_{11}^{\mathrm{i}}\right)^{-1}\mathrm{N}_{12}^{\mathrm{i}} \end{cases} \tag{38.12}$$

The single system static PPP sequential resolution based on parameter elimination can be realized using the sequential formula of Eq. (38.12).

The above method is called parameter pre-elimination. If we don't need the parameter of X_1, we can only solve X_2 after equivalent elimination steps. It doesn't need to regain the parameter of X_1.

38.3.2 The Multi-systems Combined PPP Algorithm Based on the Parameter Elimination

For the multi-systems combined positioning, the fixed parameter X_2 of observation equation of each single system can be classified into two kinds as $X_2 = [\,Y_1 \quad Y_2\,]^T$. One kind represents the overlapping parameter Y_1 of different systems, such as coordinate and troposphere parameters. The other kind represents the non-overlapping parameter Y_2 such as ambiguity parameter. Therefore N_2 and R_2 obtained from the elimination of equation (38.10) can be rewritten as a form of block matrix as follows:

$$N_2 = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},\ R_2 = \begin{bmatrix} R_{11} \\ R_{22} \end{bmatrix} \tag{38.13}$$

Expand Eq. (38.10) using Y_1 and Y_2 to matrix form as follow:

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} R_{11} \\ R_{22} \end{bmatrix} \tag{38.14}$$

The equation below can be obtained according to (38.14):

$$\begin{bmatrix} M_1 & 0 \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} R_1 \\ R_{22} \end{bmatrix} \tag{38.15}$$

where $M_1 = M_{11} - M_{12}M_{22}^{-1}M_{21}$, $R_1 = R_{11} - M_{12}M_{22}^{-1}R_{22}$. Then

$$M_1 Y_1 = R_1 \tag{38.16}$$

For GPS, GLONASS, BDS and Galileo, the same estimated parameter is Y_1, therefore the normal equation of overlapping parameters of each system can be obtained alone as follows:

$$M_1^G Y_1 = R_1^G,\ M_1^R Y_1 = R_1^R,\ M_1^C Y_1 = R_1^C,\ M_1^E Y_1 = R_1^E, \ldots, \tag{38.17}$$

where, the superscripts G, R, C and E represent GPS, GLONASS, BDS and Galileo system respectively.

Therefore the normal equation of GNSS multi-systems combined PPP solution can be obtained as follow:

$$\sum_{k=1}^{m} M_1^k Y_1 = \sum_{k=1}^{m} R_1^k \tag{38.18}$$

where m is the number of satellite navigation systems.

When Y_1 is solved, Y_2 can be regained according to Eqs. (38.14) and (38.15), then X_2 can be obtained and X_1 can be regained. At last the resolution of combined PPP is obtained. The same to Eq. (38.12), the computing formula of multi-systems combined PPP of i epoch is as follow:

$$\begin{cases} \hat{\mathrm{N}}^{\mathrm{i}}_{2(\mathrm{new})} = \hat{\mathrm{N}}^{\mathrm{i}}_{2} + \begin{bmatrix} \sum\limits_{k=1,k\neq l}^{m} M_1^k & 0 \\ 0 & 0 \end{bmatrix} \\ \hat{\mathrm{R}}^{\mathrm{i}}_{2(\mathrm{new})} = \hat{\mathrm{R}}^{\mathrm{i}}_{2} + \begin{bmatrix} \sum\limits_{k=1,k\neq l}^{m} R_1^k \\ 0 \end{bmatrix} \\ \mathrm{X}_2 = \left(\hat{\mathrm{N}}^{\mathrm{i}}_{2(\mathrm{new})}\right)^{-1} \hat{\mathrm{R}}^{\mathrm{i}}_{2(\mathrm{new})} \\ \mathrm{X}^{\mathrm{i}}_1 = \left(\mathrm{N}^{\mathrm{i}}_{11}\right)^{-1}\left(\mathrm{W}^{\mathrm{i}}_{11} - \mathrm{N}^{\mathrm{i}}_{12}\mathrm{X}_2\right) \\ \mathrm{Q}_{\mathrm{X}_2} = \left(\hat{\mathrm{N}}^{\mathrm{i}}_{2(\mathrm{new})}\right)^{-1} \\ \mathrm{Q}_{\mathrm{X}^{\mathrm{i}}_1} = \left(\mathrm{N}^{\mathrm{i}}_{11} - \mathrm{N}^{\mathrm{i}}_{12}\left(\mathrm{N}^{\mathrm{i}}_{22(\mathrm{new})}\right)^{-1}\mathrm{N}^{\mathrm{i}}_{21}\right)^{-1},\ \mathrm{N}^{\mathrm{i}}_{22(\mathrm{new})} = \hat{\mathrm{N}}^{\mathrm{i}}_{2(\mathrm{new})} + \mathrm{N}^{\mathrm{i}}_{21}\left(\mathrm{N}^{\mathrm{i}}_{11}\right)^{-1}\mathrm{N}^{\mathrm{i}}_{12} \end{cases} \tag{38.19}$$

In Eq. (38.19), $\sum\limits_{k=1,k\neq l}^{m} M_1^k$ and $\sum\limits_{k=1,k\neq l}^{m} R_1^k$ represent the sum of normal equation observation matrix of other systems except the system l itself.

38.3.3 The Adaptive Combined PPP Algorithm Based on the Posterior Variance Factor

At present, each of the three GNSS navigation systems can provide four or more visual satellites to real-time navigation positioning in global (GPS and GLONASS) or regional (BDS) area. It will provide valuable prior information when processing combined data. The independent overlapping parameters (coordinate parameter and troposphere parameter) and their precision information of single system can be obtained, which make possible to realize the adaptive combined PPP positioning in the data processing.

Therefore, the Algorithm which using the posterior variance of single system to determine the adaptive factor is proposed. Suppose the matrix form of adaptive factor of k system α_k is as follow:

$$\alpha_k = \begin{bmatrix} 1/\sigma_{x_k}^2 & 0 & 0 & 0 \\ 0 & 1/\sigma_{y_k}^2 & 0 & 0 \\ 0 & 0 & 1/\sigma_{z_k}^2 & 0 \\ 0 & 0 & 0 & 1/\sigma_{T_k}^2 \end{bmatrix} \tag{38.20}$$

where, $\sigma_{x_k}^2$, $\sigma_{y_k}^2$, $\sigma_{z_k}^2$ and $\sigma_{T_k}^2$ can be obtained by posterior variance of single system σ_0^2 multiplying by diagonal elements of corresponding parameter variance matrix Q. The specific formula can be found in references [22–24]. The Eq. (38.20) is the form of classified adaptive factors [24, 25] which are adjusted according to different parameters. In the practical work, the whole single factor can be solved by extraction of the sum of four parameters' variance and the computing formula is as follow:

$$\alpha_k = \frac{1}{\sqrt{\left(\sigma_{x_k}^2 + \sigma_{y_k}^2 + \sigma_{z_k}^2 + \sigma_{T_k}^2\right)}} \tag{38.21}$$

The single factor adjusted method as Eq. (38.21) is used in the follow-up calculating examples in this paper.

Therefore the calculating formula of the normal equation of adaptive combined PPP overlapping parameter can be obtained as follow:

$$\sum_{k=1}^{m} \alpha_k M_1^k Y_1 = \sum_{k=1}^{m} \alpha_k R_1^k \tag{38.22}$$

According to above equation, when the different GNSS systems are combined to position, the contribution of one system to the combined solution can be adaptively adjusted by the internal precision of the single system. Thus the combined solution won't be influenced when one of the systems contains blunders and the precision and stability of the combined solution can be improved.

38.4 The Algorithm Flow

The main flow of the proposed algorithm is as follows:

1. The observation equation of single system (Eq. 38.7 in current epoch can be listed independently;
2. By solving the observation equation of single system independently, the corresponding normal equation (Eq. 38.8) can be obtained;
3. Adding the prior normal equation coefficient matrix $\hat{N}_2^{i-1}$ and $\hat{R}_2^{i-1}$ of last epoch and the normal equation matrix N_2^i and R_2^i in current epoch, the

accumulative normal equation coefficient matrix $\hat{N}_2^i$ and $\hat{R}_2^i$ (The first line of Eq. 38.12) of X_2 can be obtained;

4. If it is the post solution and not in the last epoch, turn to step (10) directly, otherwise turn to step (5);
5. To do single system PPP using step (1–3), the adaptive variance factor α_k of each system can be obtained;
6. To eliminate the normal equation coefficient matrix $\hat{N}_2^i$ and $\hat{R}_2^i$ of fixed parameter obtained by procedure (3), the normal equation coefficient matrix M_1^i and R_1^i of overlapping parameter Y_1 of different systems can be obtained;
7. Weightily overlapped matrix M_1^i and R_1^i of different systems according to adaptive factor α_k, the combined solution of overlapping parameter Y_1 can be obtained;
8. According to the combined solution Y_1, the other fixed parameter Y_2 of each system is regained, thus the fixed parameter X_2 of each system can be obtained.
9. According to the fixed parameter X_2, the time-changeable parameter X_1 of each system can be regained;
10. Estimate the precision of unknown parameter;
11. Deliver the normal equation coefficient matrix $\hat{N}_2^i$ and $\hat{R}_2^i$ of fixed parameter X_2 of each single system to the next epoch; turn to step (1).

For understanding easily, the flow chart of the combined algorithm in epoch i is shown as Fig. 38.1:

In order to illustrate the character of the proposed algorithm, the brief sketch map of the new algorithm is given in Fig. 38.2.

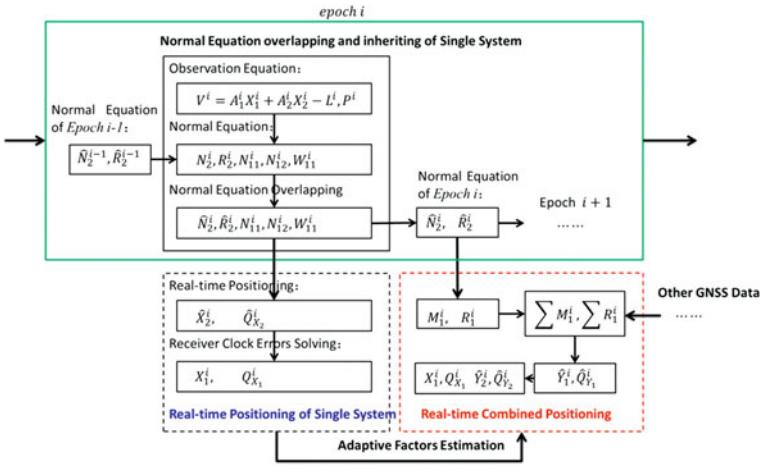


Fig. 38.1 The flow-chart of combined PPP in epoch i

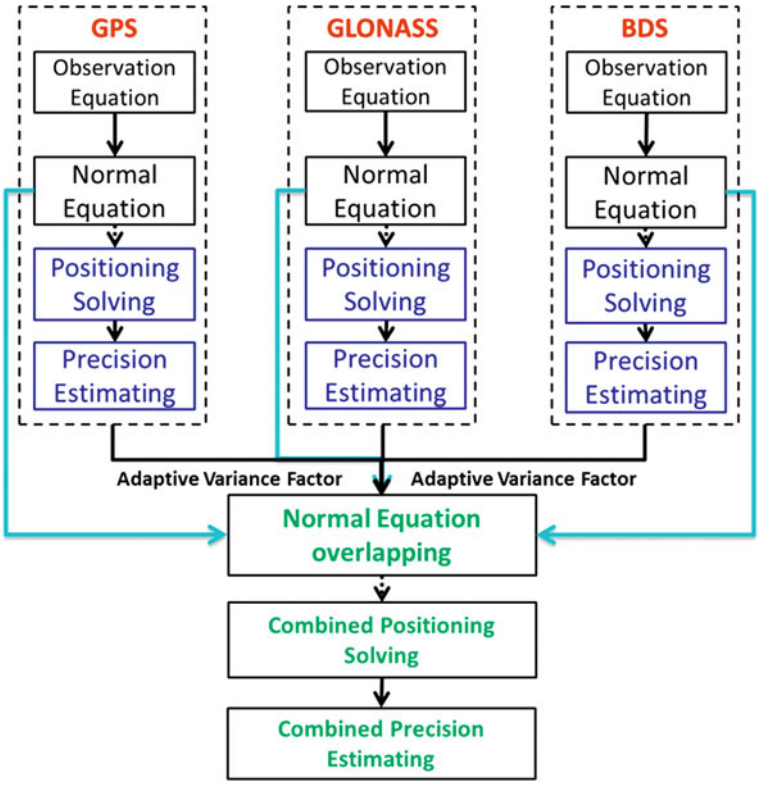


Fig. 38.2 The combined PPP algorithm based on parameter elimination

According to the solving flow and sketch map, it can be seen, compared to conventional solution, the multi-systems combined PPP algorithm based on parameter equivalent elimination has some characters as follows:

1. The contributed weight of the different systems to combined solution can be adjusted adaptively;
2. The calculating burden of the new algorithm increases linearly with the increasing of amount of combined systems. That is superior to the exponential increasing of traditional algorithm and the calculating efficiency is improved immensely;
3. If only one system or overlapping parameter are attention, the proposed algorithm can only choose a part of parameters to be solved and the calculating efficiency will be improved further;
4. For the post-processing solution, it is not necessary to solve normal equation at each epoch. The normal equation of single system can be overlapped directly and just do the combine eliminated solving at the last epoch;
5. The proposed algorithm is compatible to the single system PPP positioning software because of the fixed dimension 4×1 of overlapping parameter.

Therefore only the normal equation matrix (M1 and R1, their dimensions are 4×4 and 4×1 respectively) of overlapping parameter outputted by single system PPP software epoch by epoch are needed to realize the multi-systems combination. The change of program is small.

38.5 The Computing Examples

In order to verify the efficiency and precision of the proposed algorithm, two kinds of computing examples are designed.

38.5.1 The Comparison of Calculating Efficiency

The observations which contain 1,000 epochs of GPS, GLONASS, BDS and Galileo are simulated in there. The amount of observed satellites of each system at each epoch is 10. There is no satellite change in the observation period. Two schemes are designed to compare the calculating efficiency.

- **Scheme 1**: The multi-models combined PPP positioning based on the traditional algorithm (Sect. 38.2.2);
- **Scheme 2**: The multi-models combined PPP positioning based on the parameter elimination (Sect. 38.3.2).

Scheme 1 and scheme 2 are used to compute the combined static PPP positioning of single-system, double-systems, triple-systems and fourfold-systems. The data processing carrier is the high property collocated computer. Different numbers

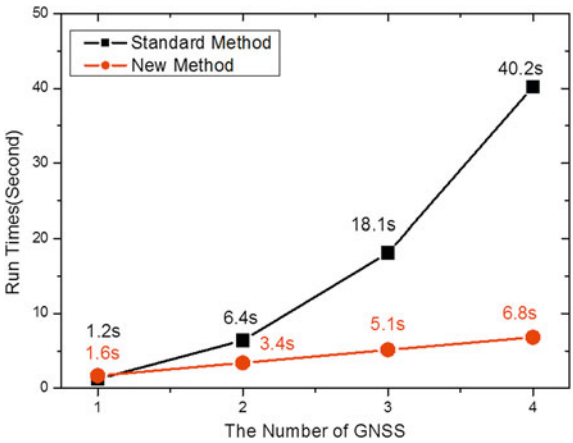


Fig. 38.3 The time information of multi-systems combined PPP

of GNSS systems are chosen to combined PPP position in the two schemes. The calculating time is given in Fig. 38.3.

From Fig. 38.3 we can find that, with the increasing of GNSS combined systems, the processing time of traditional combined algorithm increases exponentially and its calculating burden has the same type of increasing. However, the processing time of the parameter eliminated algorithm increases linearly. With the widely application of multi-models systems, the proposed algorithm has a higher calculating efficiency than the traditional algorithm.

38.5.2 The Kinematic and Static Precision Test of Combined Positioning

The measured data of GPS/GLONASS double-systems is chosen to do an analysis in order to verify the validity of the proposed combined PPP algorithm and adaptive factor. The data is downloaded from the global IGS tracking monitoring network and 24 IGS stations with GPS/GLONASS dual-frequency receivers that well-distributed on the European continent plate are chosen to be static PPP calculated. The distribution map of IGS stations is shown in Fig. 38.4. The sample interval of data is 30 s and the observed period is 24 h. The detailed processing model of static PPP can be found in Ref. [9], which will not be described here anymore.

Four schemes are designed to make an analysis.

- **Scheme 1:** GPS single system static PPP positioning;
- **Scheme 2:** GLONASS single system static PPP positioning;
- **Scheme 3:** The combined PPP positioning based on parameter elimination;

Fig. 38.4 The distribution of the chosen 24 IGS stations

- **Scheme 4:** The adaptively combined PPP positioning based on the posterior variance.

According to the "true value" which is published on IGS website, the bias of positioning results of each station in N, E and U directions is calculated. The whole RMS of each station's bias is analyzed as well. The bias of the static PPP results and its RMS value of four schemes are given in Figs. 38.5–38.8.

From Fig. 38.5–38.8 we can find that:

1. For the single-day solution, the single system PPP positioning (Figs. 38.5 and 38.6) and double-systems combined PPP positioning (Figs. 38.7 and 38.8) can achieve a horizontal precision of mm and a vertical precision of about 1 cm, which show the good observed condition of IGS station and the character of high-precision of the current PPP technology;

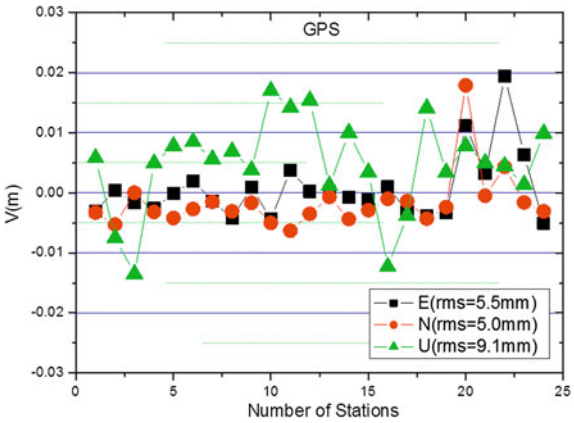


Fig. 38.5 The bias of positioning result and RMS of scheme1

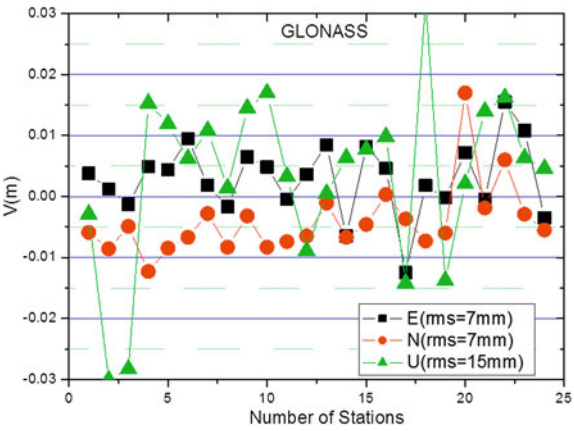


Fig. 38.6 The bias of positioning result and RMS of scheme 2

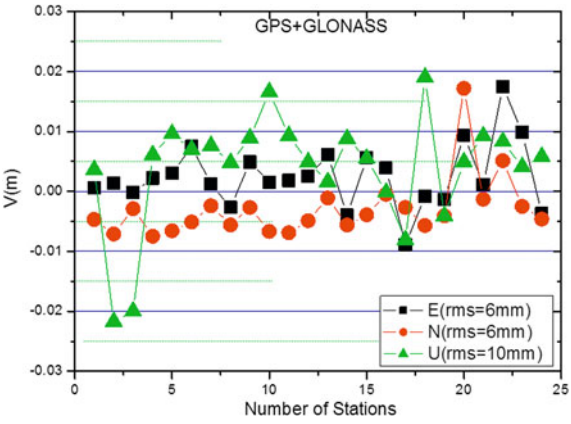


Fig. 38.7 The bias of positioning result and RMS of scheme 3

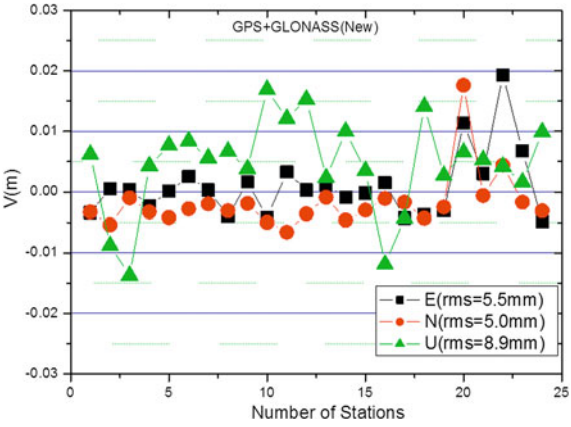


Fig. 38.8 The bias of positioning result and RMS of scheme 4

2. For the single system positioning, the precision of GLONASS positioning (Fig. 38.6) is a bit lower than GPS positioning (Fig. 38.5) and the precision difference in height direction is bigger than in horizontal direction. However, the difference between them has a magnitude of mm. It shows the precision of GLONASS positioning is approximate to GPS and it has gotten rid of the condition that had a precision one magnitude lower than GPS in early age. This is a good basis for the realization of the valid combination of two systems;
3. Since scheme 3 cannot distribute the weight of GPS and GLONASS observations rationally and only overlap the normal equation with the equal weight, the precision of its combined positioning result is between the two single systems. While the problem is solved and the optimal precision is obtained by adaptively combined positioning by using posterior variance (scheme 4). What has to be mentioned here is that since the single system static PPP positioning can

achieve a precision of cm at present, the contribution of combined positioning in good observed condition is not significant.

Moreover, in order to verify the validity and precision of the proposed algorithm in kinematic PPP positioning, a high-frequency station (ATAC) with GPS/GLONASS dual-models receiver has been chosen to make kinematic PPP calculation. The sample interval of data is 1 s and the whole observed period is 5 h. The detailed processing model of kinematic PPP can be found in Ref. [26].

According to the "true value" published on IGS website, the bias of kinematic positioning result of the station in N, E and U directions (the 30 min' convergent trajectory has been eliminated) is calculated. The corresponding RMS of the station's bias is analyzed as well. The bias of the kinematic PPP results and its RMS value are given in Figs. 38.9–38.12.

From Figs. 38.9–38.12 we can find that:

1. Similar to the static positioning result, the precision of GLONASS kinematic positioning (Fig. 38.10) is a bit lower than GPS positioning (Fig. 38.9) and the precision difference in height direction is bigger than in horizontal direction;
2. The kinematic positioning precision (Figs. 38.11 and 38.12) of both equal-weight combined PPP (scheme 3) and adaptively combined PPP (scheme 4) is higher than the single system PPP positioning results (Figs. 38.9 and 38.10), which shows the great contribution of multi-systems combined positioning to the kinematic positioning precision improvement;
3. In the good observed condition, the positioning precision of scheme 3 and scheme 4 is mainly equal.

In order to analyze the validity of scheme 4 when there are blunders in observations, the first satellite of each epoch has been added a phase error of 1 m from 10,000th epoch to 12,000th epoch of the kinematic station. The observation

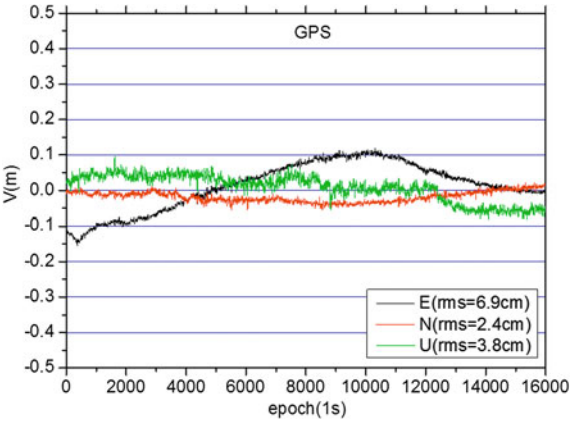


Fig. 38.9 The bias of kinematic positioning and RMS of scheme 1

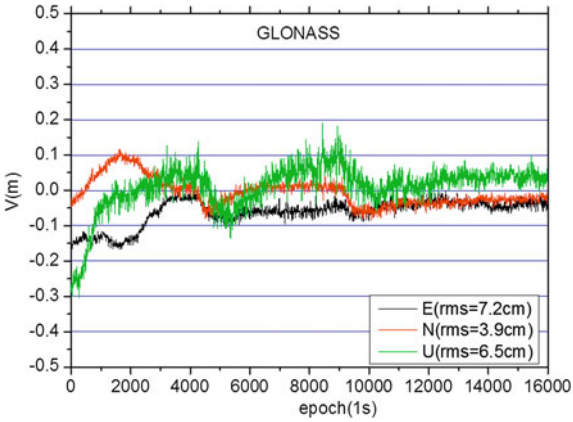


Fig. 38.10 The bias of kinematic positioning and RMS of scheme2

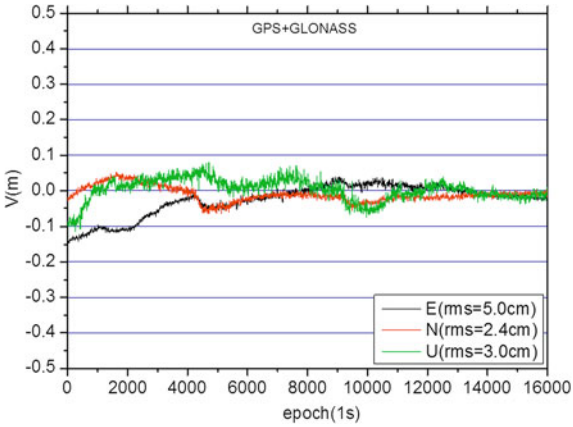


Fig. 38.11 The bias of kinematic positioning and RMS of scheme 3

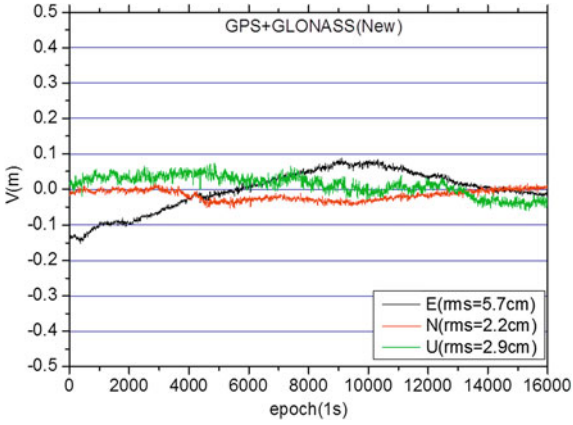


Fig. 38.12 The bias of kinematic positioning and RMS of scheme 4

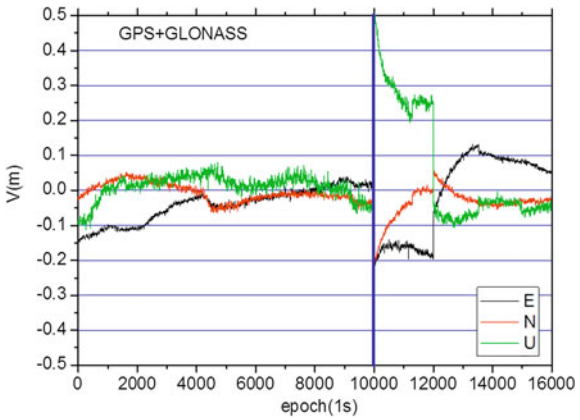


Fig. 38.13 The bias of kinematic positioning and RMS of scheme 3 (including blunders)

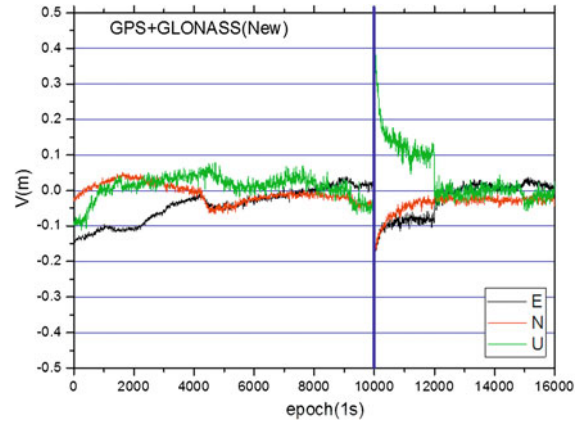


Fig. 38.14 The bias of kinematic positioning and RMS of scheme 4 (including blunders)

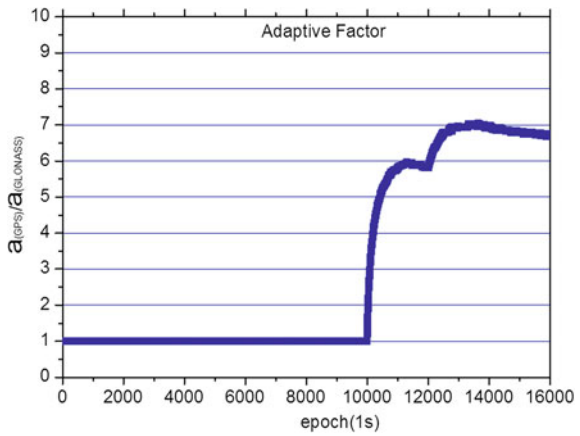


Fig. 38.15 The variation of the adaptive factor ratio of scheme 4

data after 10,000th epoch is calculated by scheme 3 and scheme 4 to make a comparison, and the positioning results are shown as Figs. 38.13 and 38.14.

The variation of the adaptive factor ratio of scheme 4 is shown as Fig. 38.15.

From Figs. 38.13–38.15 we can find that, when there exit blunders in the observations of single system, the positioning result of scheme 3 is seriously influenced while it only has a bit influence on scheme 4. The posterior variance of single system will increase gradually when there is observed error in the single system. Then the adaptive factor in scheme 4 will adjust adaptively, that the combined weight of different systems will be distributed again and the contribution of the normal equation of the system including blunders are decreased and the blunders are restrained.

It should be mentioned that although this adaptive method can distribute the combined weight by posterior variance, the blunders cannot be detected at once in case the overlapping parameter was obtained by the whole sequential solution of several epochs. Therefore, except the proposed method, the adaptive processing of blunder needs more flexible method which is also the study point of the authors later.

38.6 Conclusions and suggestions

The traditional GNSS combined PPP algorithm cannot meet the demand of the rapid resolution of high-precision real-time high-frequency positioning. Therefore a combined sequential PPP algorithm based on parameter equivalent elimination principle is proposed. First the whole combined observations of multi-models PPP is divided into each single system to be solved independently and paralleled. Second the normal equation matrixes of overlapped parameters which belong to the different systems have been accumulated. Third the combined resolution can be obtained directly by solving the accumulated normal equation matrix. The new algorithm improves the calculating efficiency of the GNSS combined Positioning immensely. In the meanwhile, a combined positioning method which uses posterior variance factors to adaptively adjust the contributed weight relations of different GNSS systems is also proposed. At last, the availability and precision of the new algorithm are verified by analysis of different algorithms' calculating time and the PPP positioning precision.

The parameter equivalent elimination algorithm can be applied into the researches on GNSS precise satellite clock resolution, the un-different baseline network adjustment and GNSS time offset monitoring, and it also can improve the calculating efficiency effectively.

Acknowledgments This study was supported by National Natural Science Foundation of China (Grant No. 41020144004, 40902081, 41104019 and 41104022), the Special Fund for Basic Scientific Research of Central Colleges (Grant No. CHD2010ZY001, No. CHD2011ZY023, Chang'an University).

References

1. Yang Y (2010) Progress, contribution and challenges of compass/beidou satellite navigation system. Acta Geodaetica et Cartographica Sinica 39(1):1–6
2. Yang Y, Li J, Xu J et al (2011) Contribution of the compass satellite navigation system to global PNT users. Chinese Sci Bull 2011:56. doi:10.1007/s11434-011-4627-4
3. Ge M, Chen J, Dousa J, Gendt G, Wickert J (2012) A computationally efficient approach for estimating high-rate satellite clock corrections in real-time. GPS Solutions. doi 10.1007/s10291-011-0206-z
4. Ge M, Gendt G, Dick G, Zhang FP (2005) Improving carrier-phase ambiguity resolution in global GPS network solutions. J Geodesy 79(1–3):103–110
5. Ge M, Gendt G, Rothacher M, Shi C, Liu J (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations. J Geodesy 82(7):389–399
6. Ge M, Chen J, Dousa J, et al. (2011) A computationally efficient approach for estimating high-ratesatellite clock corrections in realtime. GPS Solutions. Doi 10.1007/s10291-011-0206-z
7. Zhang B, Ou J, Yuan Y, Zhong S (2010) Precise point positioning (PPP) algorithm based on original dual-frequency GPS code and carrier-phase observations and its application. Acta Geodaetica et Cartographica Sinica 39(5):478–483
8. Ye SR (2002) GPS precise point positioning using un-differenced phase observation. Wuhan University, Wuhan
9. Huang GW (2008) Research on GPS precise point positioning and adjustment of high precision GPS baseline network and the realization of its software. Chang'an University, Xian
10. Wang ZJ (2012) Performance analysis of combined GPS/GLONASS precise point positioning. J Geodesy Geodyn 2(32):105–109
11. Cai CS et al (2011) Performance analysis of combined GPS/GLONASS precise point positioning. J Geodesy Geodyn 3:85–89
12. Zhang XH, Guo F, Li XX, et al. (2011) Study on precise point positioning based on combined GPS and GLONASS. Geomatics Inf Sci Wuhan Univ 35(1):9–12
13. Melgard T, Jong KD, Lachapelle G, et al. (2011) Interchangeable integration of GPS and GLONASS by using a common system clock in PPP. In: ION GNSS 2011, Session F3, Portland, OR, 20–23 Sept 2011
14. Jokinen A, Feng SJ, Milner C, et al. (2011). Precise point positioning and integrity monitoring with GPS and GLONASS[C], [Online]. Available: http://www.rin.org.uk/Uploadedpdfs/ConferenceProceedings/
15. Cai CS (2009) Precise point positioning using dual-frequency GPS and GLONASS measurements. University of Calgary, Master Paper
16. Melgard T, Vigen E, Orpen O (2009). Advantages of combined GPS and GLONASS PPP—experiences based on G2, a new service from Fugro. In: 13th IAIN world congress, 27–30 Oct, Stockholm, pp 1–7
17. Russell D (2010) Improving accuracy and redundancy with GPS and GLONASS PPP. In: Dynamic positioning conference 12–13 Oct, pp 1–12
18. Xu GC (2003) gps theory, algorithms and applications. Springer, Berlin
19. Xu GC (2002) GPS data processing with equivalent observation equations. GPS Solutions 6:28–33
20. Wang QX, Xu TH, Xu GC (2011) Adaptively changing reference station algorithm and its application in GPS long range airborne kinematic relative positioning. Acta Geodaetica et Cartographica Sinica 4(40):429–434
21. Yang YX, Zeng AM (2008) Fusion modes of various geodetic observations and their analysis . Geomatics Inf Sci Wuhan Univ 33(8):771–774
22. Yang YX, Song LJ, Xu TH (2002) Robust parameter estimation for geodetic correlated observations. Acta Geodaetica et Cartographica Sinica, 31(2):95–99

23. Yang YX, Song LJ, Xu TH (2002) Robust estimator for correlated observations based on bifactor equivalent weights. J Geodesy 76(6–7):353–358
24. Yang YX, Cui XQ (2008) Adaptively robust filter with multi adaptive factors. Survey Rev 40(309):260–270
25. Huang GW, Yang YX, Zhang Q (2011) Estimate and predict satellite clock error using adaptively robust sequential adjustment with classified adaptive factors based on opening windows. Acta Geodaetica et Cartographica Sinica 40(1):15–21
26. Huang GW, Tu R, Zhang Q et al (2010) A new real-time kinematic precise point positioning algorithm based on reference-station corrections. J Geodesy Geodyn 36(6):135–139

Author Biography

Huang Guanwen received a Bachelor and M.S. degrees in Surveying Engineering from Chang'an University of China in 2005 and 2008. He begins his PhD study in 2008 in the same university, and currently works in the Institute of GFZ in Germany as a PhD student from 2011 to 2012. His research activities include Precise Point Positioning theory, real-time satellite clock model and their application in Global Navigation Satellite System.

Chapter 39
Impact of Ionospheric Correction on Single-Frequency GNSS Positioning

Ningbo Wang, Yunbin Yuan, Zishen Li and Xingliang Huo

Abstract The ionospheric delay is one of the most important error sources affecting the accuracy of Global Navigation Satellite Systems (GNSS), especially for single-frequency users. It is important to investigate various methods to mitigate the ionospheric effects on GNSS single-point single-epoch (SPSE) positioning when using single-frequency receivers. Based on the GPS data gathered at 5 International GNSS Service (IGS) stations at the mid-low latitude regions in China, we accessed the performances of different ionospheric corrections in terms of ionospheric Total Electron Content (TEC) correction accuracy and SPSE positioning during periods of high (2002) and low (2009) solar activity respectively. Meanwhile, the impacts of ionospheric corrections on GNSS positioning (especially on horizontal accuracy) were analyzed, and the relationship between TEC and positioning residual errors was also discussed. The results show that all ionospheric corrections can obviously improve the vertical accuracy, while the horizontal improvement is less than the vertical one, and even worse, the horizontal error may become larger with ionospheric corrections. Furthermore, the north error plays an important role in limiting the horizontal accuracy. If north error was not properly corrected with ionospheric models, it would worsen the horizontal positioning performance.

Keywords Single-frequency GNSS positioning · Ionospheric correction · Total Electron Content (TEC) · Percentage of vTEC RMS correction (PTRC)

N. Wang (✉) · Y. Yuan · Z. Li · X. Huo
State Key Laboratory of Geodesy and Earth's Dynamics, Institute of Geodesy and Geophysics, Wuhan, China
e-mail: wnbigg@asch.whigg.ac.cn

N. Wang
University of Chinese Academy of Sciences, Beijing, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_39,

39.1 Introduction

The ionospheric delay is considered as one of the most critical sources of error affecting the accuracy of GNSS. The amount of ionospheric delay is dependent on GNSS signal frequency and Total Electron Content (TEC) along the signal propagation path from the satellite to the receiver. For dual-frequency users, the ionospheric delay can be eliminated by using the geometry-free linear combination, while it has to be mitigated with different ionospheric models for real-time single-frequency (SF) navigation and positioning users [1]. The Global Positioning system (GPS) uses the well-known Klobuchar model to correct the ionospheric error and the Galileo system uses NeQuick model to mitigate the effect of ionospheric delay [2, 3].

A vast number of studies have been conducted on the evaluation and comparison of different ionospheric models for ionospheric delay mitigation, while most of these studies focus on ionospheric TEC performance [4–8]. Compared with the performance of these ionospheric corrections expressed in terms of TEC or ionospheric delay, SF navigation users are more interested in ionospheric impact on positioning accuracy, which has been studied in recent years [9–12]. For instance, the bias propagation of ionospheric delay into the final SPSE positioning has been analyzed and it is reported that the TEC horizontal gradient is the main contributor of horizontal positioning errors [9]. The performance of Klobuchar and NeQuick algorithms at mid-latitudes has also been accessed and it is stated that NeQuick model improves more the vertical accuracy than the horizontal one [10]. Another interesting study carried by Chen et al. [11] shows a large decrease of positioning 3D RMS (7.8–35.3 %) with COMPASS regional ionospheric model compared to Klobuchar model in northern hemisphere. However, there is no overall assessments of different ionospheric corrections at the mid-low latitudes in China, where is seriously affected by the well-known equatorial ionization anomaly (EIA).

With regard to it, the impact of several single-frequency ionospheric corrections on GNSS positioning in China mid-low latitudes was investigated in this paper. Based on the GPS vertical TEC (vTEC) data derived at 5 IGS stations during periods of high (2002) and low (2009) solar activity, firstly the effectiveness of different ionospheric models was accessed in terms of TEC correction accuracy. Secondly the positioning performances by using these ionospheric corrections were presented, which had a clear meaning for SF navigation users. Finally the impact of ionospheric correction on positioning in particular for horizontal components was analyzed, and the relationship between TEC performance and positioning accuracy was discussed. By evaluating the performance of different single-frequency ionospheric correction algorithms, it will provide some meaningful help for the establishment and assessment of the future COMPASS global broadcast ionospheric model.

39.2 Data Processing and Analysis Methods

39.2.1 Ionospheric TEC Measurements

Obtaining ionospheric TEC measurement using dual-frequency GPS observations, has been extensively discussed [13]. In order to remove the ambiguity of the carrier phase and to reduce the code noise and multipath, the strategy of "phase-smoothed code measurements" is employed. The GPS-derived ionospheric TEC measurement can be expressed as Eq. (39.1).

$$P = 40.3 \cdot \left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right) \cdot TEC + c \cdot (\Delta b^s + \Delta b_r) \tag{39.1}$$

where P denotes TEC and Differential Code Biases (DCB) measurement obtained with the processing of "carrier phases aligned to the code" for individual satellites in a consecutive arc (in m);

f_1 and f_2 are the two frequencies of the GPS carrier phase (in Hz);

$\Delta b^s = b_1^s - b_2^s$ and $\Delta b^r = b_{1,r} - b_{2,r}$ are DCB of satellite and receiver respectively (in s), where b_1^s is the satellite hardware on frequency f_1 (b_2^s on frequency f_2) and $b_{1,r}$ is the receiver hardware on frequency f_1 ($b_{2,r}$ on frequency f_2). They are removed directly with the CODE-released values in the following section [14].

c denotes the speed of light (in m/s).

39.2.2 Accuracy Assessment Techniques

The performance of different ionospheric models was investigated by using the GPS data gathered at 5 IGS stations (WUHN, SHAO, LHAS, KUNM, and TWTF) located at mid-low latitudes in China. The distribution of these selected IGS stations is shown in Fig. 39.1. We also gave the data availability levels for the

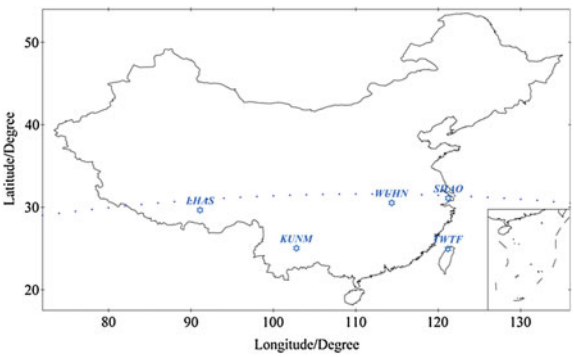


Fig. 39.1 Location of the selected 5 IGS stations in China's mid-low latitudes (The *blue dotted line* represents the geomagnetic latitude +20°)

selected year: it provides 91 % of RINEX data availability for 2002 and 97 % for 2009 respectively due to the lack of GPS raw data at some IGS stations.

In order to evaluate the performance of different ionospheric models, two kinds of analysis methods in terms of ionospheric TEC correction accuracy and positioning accuracy were used. Based on the vTEC derived from each ionospheric model and that from actual GPS data, two vTEC statistical indices were presented to analyze the ionospheric TEC correction accuracy, as shown in Eqs. (39.2) and (39.3).

$$RMS_{TEC} = \sqrt{\left\langle \left(vTEC_{\mathrm{model}} - vTEC_{meas}\right)^2 \right\rangle} \tag{39.2}$$

$$PTRC = 1 - \frac{RMS_{TEC}}{\langle vTEC_{meas} \rangle} \cdot 100\% \tag{39.3}$$

where $< >$ denotes the mean value on a given time frame;

RMS_{TEC} is the vTEC root mean square (RMS) calculated with modeled and measured vTEC values (in TECu, 1TECu equals 10^{16} electrons per m^{-2} and corresponds to a range delay of 16 cm at GPS L1 frequency);

PTRC is the percentage of vTEC RMS correction, which is often used to quantify the performance of different ionospheric models (in %) [2, 5, 6, 10].

The impact of ionospheric corrections on positioning accuracy was discussed based on the GNSS SPSE positioning. The receiver positions were calculated using GPS C/A pseudorange measurements, broadcast ephemeris, Hopfield tropospheric model, the Niell mapping function as well as several SF ionospheric correction algorithms. Different ionospheric models include:

- NoIon—Without any ionospheric corrections, i.e., the uncorrected case;
- Klobuchar—GPS broadcast ionospheric model with 8 coefficients designed by Klobuhcar;
- Ref-Klobuhcar—Klobuchar-style ionospheric coefficients (4 alphas and 4 betas) provided by the Center for Orbit Determination in Europe (CODE);
- CODESH—Global ionospheric maps (GIM) using a spherical harmonics expansion up to degree and order 15 in Bernese format produced by CODE.

The positioning results of different ionospheric corrections were presented in terms of horizontal and vertical RMS values, by the comparison between the real position of the selected sites and the ones calculated with and without each ionospheric model. In order to analyze the positioning improvement with different ionospheric corrections, ratio value is defined as Eq. (39.4).

$$Ratio = \frac{RMS_{unc} - RMS_{\mathrm{mod}}}{RMS_{unc}} \cdot 100\% \tag{39.4}$$

where RMS_{mod} is the positioning error obtained with each ionospheric model and RMS_{unc} is the uncorrected positioning error. Positive ratio value shows improvement in positioning accuracy while negative value indicates that the positioning result calculated with ionospheric model is worse than the uncorrected case.

39.3 Ionospheric Correction Performances

39.3.1 Ionospheric TEC Correction Accuracy

From all the available GPS vTEC data and corresponding modeled values with each ionospheric correction at the 5 selected sites during the year 2002 and 2009, we calculated the percentage of vTEC RMS correction.

According to vTEC statistics (see Fig. 39.2), it can be found that Ref-Klobuchar and CODESH model improve 9.1 and 28.3 % of the PTRC respectively compared to Klobuchar model (57.5 %) in 2002. Although Klobuchar algorithm meets the declared goal of 50 % vTEC RMS correction in China mid-low latitudes, it underestimates vTEC by 13.4 TECu compared to the mean measured values. Ref-Klobuchar and CODESH models present vTEC overestimation (4.4 and 1.5 TECu) inversely. During the year 2009, Klobuchar algorithm only corrects 46.8 % RMS of the ionospheric delay, while the ionospheric vTEC correction accuracy increases much more with Ref-Klobuchar and CODESH models (9.4 and 32.4 %). In addition, vTEC overestimations are observed with all the ionospheric corrections (3.7, 3.2 and 0.2 TECu) in 2009.

Yearly vTEC statistical results also show us the influence of solar activity: the PTRC obtained with each ionospheric model is higher in 2002 than that of 2009. By analyzing the measured vTEC during periods of high and low solar activity, it can be obviously found that the mean measured vTEC in 2009 is much smaller compared to the value in 2002 (10.5 versus 45.0 TECu). Thus a small TEC RMS value during solar minimum will induce a great relative RMS error. The different vTEC performances during solar maximum and minimum indicate us that it is hard to evaluate ionospheric TEC correction accuracy by using individual statistical indice. Therefore, the PTRC (relative error) and TEC RMS (absolute error) are suggested to be employed together for the assessment of ionospheric TEC correction accuracy in practice (see Fig. 39.3).

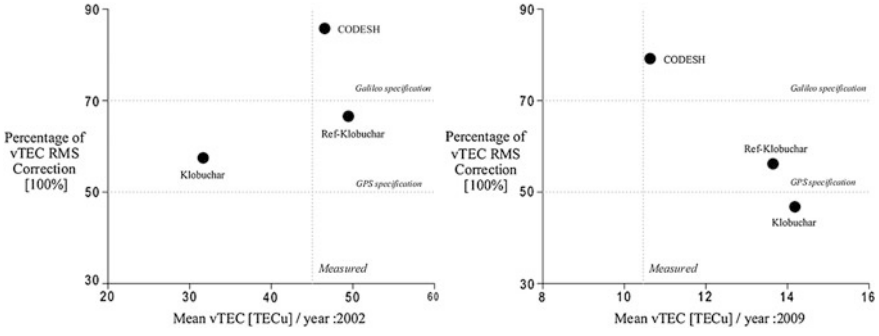


Fig. 39.2 The PTRC obtained with different ionospheric corrections in 2002 (*left plot*) and 2009 (*right plot*)

39.3.2 Positioning Accuracy

Based on the analysis method described in Sect. 39.2.2, we performed the horizontal and vertical positioning RMS results of each ionospheric model.

Figure 39.4 shows that the positioning accuracy can be improved with all the ionospheric corrections compared to the uncorrected case, while these corrections improve more the vertical accuracy than the horizontal one. The vertical error in 2002 (9.9 m) decreases obviously with Klobuchar and Ref-Klobuchar algorithms (−32 and −31 %) and even more with CODESH approach (−42 %). During the year 2009, significant decrease is also presented in vertical component for Klobuchar, Ref-Klobuchar and CODESH corrections (−39, −42 and −44 % respectively). However, it should not be ignored that the horizontal error in 2009 (1.8 m) drops much less for Ref-Klobuchar and CODESH methods (−13 and −26 %) and barely with Klobuchar model (−1 %). For the year 2002, CODESH model decreases the horizontal error (4.8 m) to 3.7 m (decreasing by 22 %), while the Klobuchar and Ref-Klobuchar algorithms improve the horizontal error to 5.8 m (increasing by 21 %), which is even larger than that of the uncorrected case.

Figure 39.5 gives the percent occurrence of the days of Klobuchar, Ref-Klobuchar and CODESH model with the smallest horizontal and vertical positioning RMS [15]. It can be found that CODESH performs best in improving the horizontal and vertical accuracy. In 2002, CODESH has the smallest daily horizontal (resp. vertical) RMS in at least 69 % and at most 94 % (resp. 75 and 94 %) of the observed days across all the stations, and the performance of Ref-Klobuchar is generally better than Klobuchar in both horizontal and vertical directions. In 2009, CODESH can provide at least 66 % and at most 85 % (resp. 61 and 89 %) of the occurrence with the smallest horizontal (resp. vertical) RMS at the 5 selected sites. Meanwhile, Klobuchar shows the best horizontal accuracy in at least 6 % and at most 28 % of the observed days and Ref-Klobuchar presents the best vertical accuracy in at least 7 % and at most 20 % of the occurrence, which implies that Klobuchar (resp. Ref-Klobuchar) improve more the horizontal (resp. vertical) accuracy compared with Ref-Klobuchar (resp. Klobuchar) in 2009.

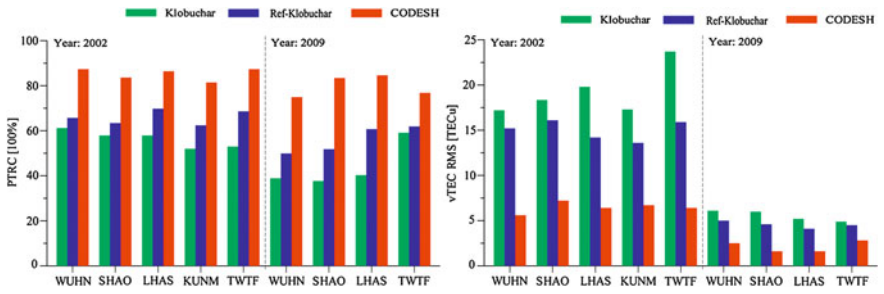


Fig. 39.3 PTRC (*left plot*) and TEC RMS (*right plot*) obtained with different ionospheric corrections at the 5 selected sites. The missing of the KUNM station in 2009 is caused by the lack of corresponding CODE-released receiver DCB values

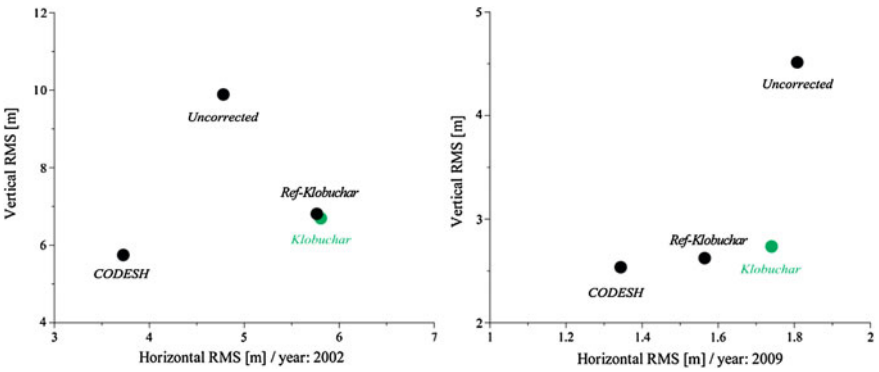


Fig. 39.4 Positioning accuracy obtained with and without the ionospheric corrections in 2002 (*left plot*) and 2009 (*right plot*)

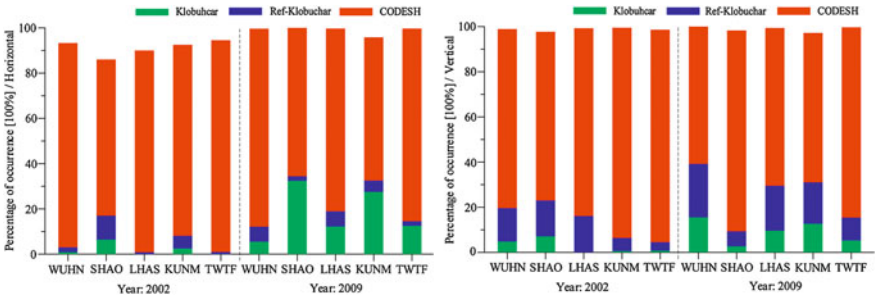


Fig. 39.5 Percent occurrence of the days with the smallest horizontal (*left plot*) and vertical (*right plot*) positioning RMS in 2002 and 2009

The observed results show horizontal positioning accuracy improving much less compared to the vertical one for different ionospheric models. Hence it indicates to investigate the impact of different ionospheric models on positioning (in particular for horizontal error) in the following discussion.

39.4 Ionospheric Influence on Positioning Accuracy

39.4.1 The Impact of Ionospheric Correction on Vertical Accuracy

In Sect. 39.3.2, the performance of different ionospheric models in terms of positioning during solar maximum and solar minimum was accessed and significant decrease in vertical error was found with all ionospheric corrections, in

particular for CODESH model. Figure 39.6 presents the vertical and horizontal ratio of each ionospheric correction across all the stations in the year 2002. The values demonstrate the improvement made by each ionospheric model compared to the uncorrected case. It is clearly shown that vertical accuracy increases by at least 15 % with Ref-Klobuchar at the station KUNM and at most 53 % with CODESH at the station WUHN, and CODESH presents the best vertical improvement (increasing by 26–53 %) at the 5 stations (see Fig. 39.6, left plot). Horizontal accuracy obtained with Klobuchar and Ref-Klobuchar model is worse than the uncorrected case across the 5 selected sites except for the station SHAO. At the station WUHN and LHAS, horizontal accuracy even decrease by 31 and 34 % with Klobuchar and Ref-Klobuchar corrections respectively (see Fig. 39.6, right plot). On the contrary, CODESH increases the horizontal accuracy by 16–28 %, which is evidently better than the Klobuchar and Ref-Klobuchar models.

As we know, vertical accuracy is influenced by the ionospheric delays above the receiver. However, the impact of ionosphere on horizontal accuracy is presented in north–south and east–west directions, which is more complicated than that of the vertical one. Hence we can expect remarkable improvement in vertical accuracy with different ionospheric corrections.

Figure 39.7 gives the distribution of vertical positioning errors with each ionospheric correction at the station of WUHN, and an obvious vertical mean error (9.8 m) toward the up direction for the uncorrected case can be found. Klobuchar algorithm partially mitigates the vertical positioning errors (4.2 m on average), and Ref-Klobuchar approach performs much better, decreasing vertical mean error below 0.3 m. Although CODESH model obtains a larger mean value (1.5 m) than the Ref-Klobuchar model, it provides a smoother and narrower vertical error distribution with a lower standard deviation (5.0 m). Hence, it can be concluded that the vertical uncorrected positioning errors are best corrected with CODESH and better with Ref-Klobuchar than with Klobuchar algorithm, which performs worst among these three corrections.

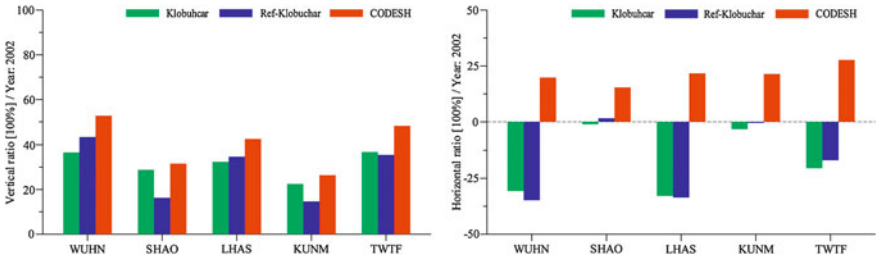


Fig. 39.6 Vertical (*left plot*) and horizontal (*right plot*) ratio obtained with different ionospheric corrections at the 5 selected sites in 2002

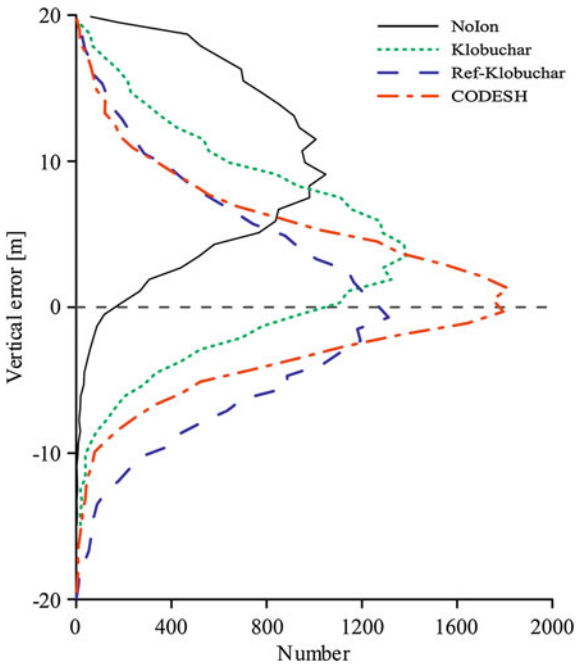


Fig. 39.7 Distributions of the vertical positioning errors with and without ionospheric corrections at the station of WUHN in 2002

39.4.2 The Impact of Ionospheric Correction on Horizontal Accuracy

As described before, good vTEC performances are obtained with all ionospheric corrections. The vertical positioning errors are also significantly mitigated, while ionospheric correction increases more the vertical accuracy than the horizontal one. In the current section, we focused on the impact of ionospheric correction on horizontal accuracy. As shown in Fig. 39.8, the distribution of horizontal uncorrected positioning errors at the station of WUHN is characterized by a large dispersion along the north direction than along the east one. To interpret the observed error distribution, we plotted the diurnal variation of the north and east uncorrected errors at WUHN station in 2002 (see Fig. 39.9). It can be clearly found that the north error is almost positive along the time axis, while the east error shows small values which are negative in the hours near sunrise and positive near sunset. By averaging the north and east errors, horizontal bias towards the north direction can be obtained.

The north positioning errors are generally positive and cannot be totally removed by ionospheric corrections. Figure 39.10 illustrates the relation between north error and horizontal error. It can be clearly seen that horizontal error increases with the north error. Particularly, larger north errors are observed with Klobuchar and Ref-Klobuchar model compared with the uncorrected case, which would result in even worse horizontal accuracy. Hence, it can be concluded that

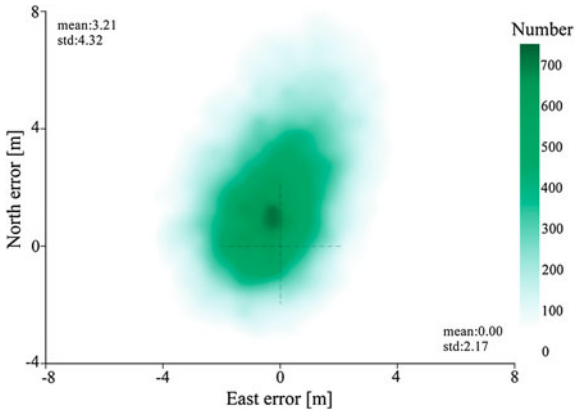


Fig. 39.8 Distributions of the horizontal uncorrected positioning errors at WUHN in 2002

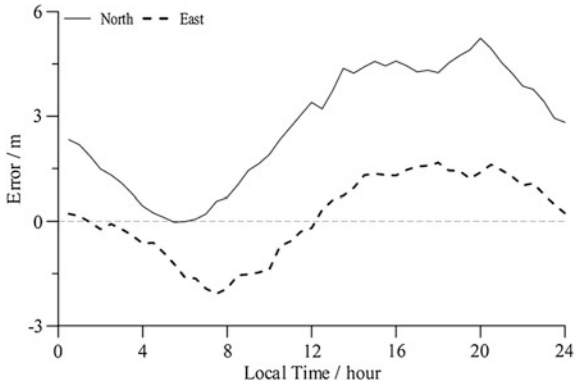


Fig. 39.9 Local-time variability of the north (*solid line*) and east (*dashed line*) uncorrected positioning errors at WUHN in 2002

the north error plays an important role in limiting the horizontal positioning accuracy. If north error was not sufficiently corrected, it would worsen the horizontal positioning performance (see Fig. 39.6, right plot).

The study conducted by Mohino shows that horizontal components are related to the horizontal TEC gradients: a southern horizontal TEC gradient (i.e., the bias with higher TEC values in the south) would contribute a positive north error in northern hemisphere, and a western horizontal TEC gradient would render a positive east error [9]. Accordingly, positive north errors imply potential southward TEC gradients at WUHN station. In the hours near sunrise, high ionospheric delays for the eastern satellites can be observed compared to the western ones because the sun is rising in the east while it is still dark in the west. Hence we can expect negative east errors during sunrise and the opposite results during sunset (see Fig. 39.9).

Figure 39.11 shows the monthly horizontal accuracy at the station WUHN in 2002. Maximal horizontal errors can be found around the equinoxes. CODESH model can improve the horizontal accuracy effectively, while Klobuchar and Ref-Klobuchar corrections make the horizontal accuracy even worse than the uncorrected case across the whole year. Considering the key effect of north bias on

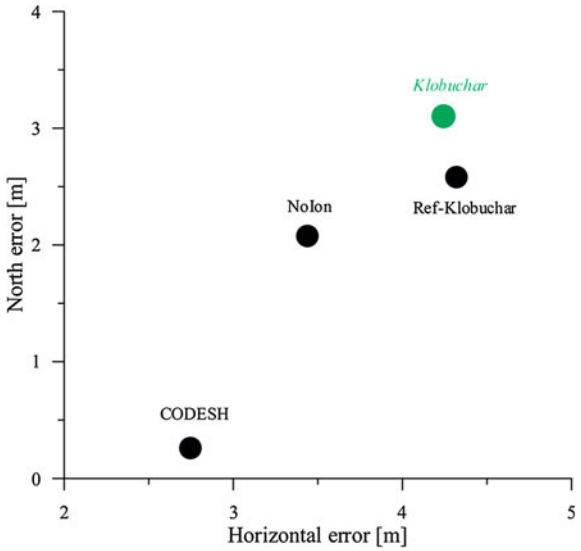


Fig. 39.10 Relation between the north error and horizontal error obtained with and without ionospheric corrections at the station of WUHN in 2002

horizontal accuracy, we analyzed the performances of Klobuchar and CODESH model on correcting the north–south TEC gradients between day of year (DOY) 244 and 273 (corresponding to Sep., the worst month) in 2002 at WUHN station.

In Fig. 39.12, the left plot presents the distributions of horizontal positioning errors with Klobuchar model, and the right plot presents the distributions of sTEC mean differences between GPS-derived sTEC and modeled values for 30° elevation and 45° azimuth bins. The circle surface represents the number of ionospheric pierce points (IPP) above the station in each bin and different colors correspond to the values of mean differences. It can be seen that horizontal positioning errors corrected with Klobuchar model display an obvious offset toward the north direction. Meanwhile, Klobuchar model underestimates the sTEC values in the south direction and shows better sTEC performance in the north direction. Figure 39.13 shows the relation between the distributions of horizontal errors and sTEC residual errors for CODESH model. It can be found that the north positioning errors are successfully corrected with the CODESH algorithm. And at the same time, CODESH model presents good sTEC performances in both of the north and south directions.

Preliminary results show that Klobuchar algorithm underestimates the slant TEC (sTEC) on the south direction, while it provides effective modeled sTEC with respect to measured values on the north direction. Horizontal gradients are partially removed by Klobuchar model, causing apparent horizontal positioning errors toward the north axis (see Fig. 39.12, left plot) and a large horizontal positioning error (12.2 m in the 95th percentile). On the contrary, CODESH model shows good sTEC performance on both of the north and south directions, and the north–south gradients mismodelling are also eliminated properly with respect to Klobuchar model. Hence, the distributions of horizontal positioning errors corrected

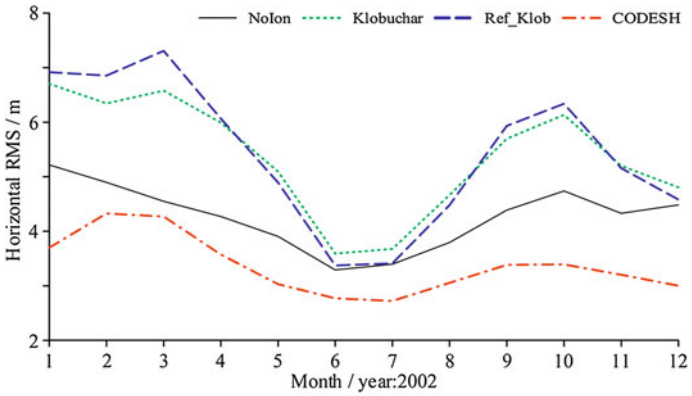


Fig. 39.11 Monthly variability of horizontal RMS obtained with and without ionospheric corrections at WUHN

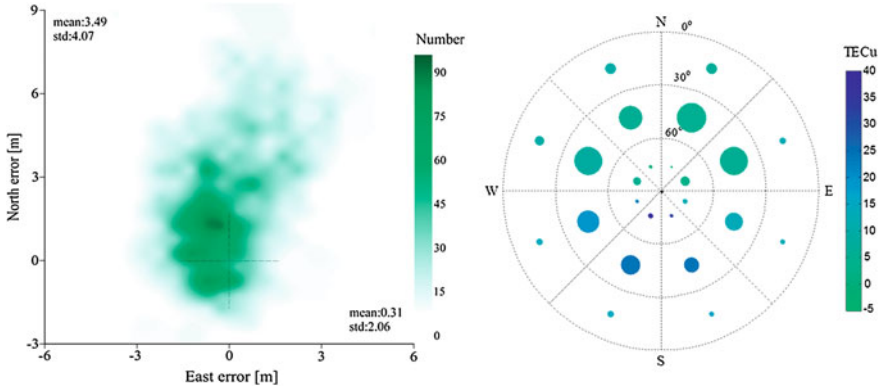


Fig. 39.12 Relation between the distributions of horizontal errors (*left plot*) and sTEC residual errors (*right plot*) for Klobuchar model at WUHN

with CODESH model are more centered (see Fig. 39.13, left plot), resulting in a better horizontal positioning performance (6.8 m in the 95th percentile). Similar results are also obtained with NeQuick model at Brussels by Bidaine [10].

39.4.3 Relation Between TEC Correction Performance and Positioning Accuracy

People frequently tend to believe that there exists an obvious relation between ionospheric vTEC performances and the corresponding positioning errors, i.e., a small positioning residual error is related to a little relative vTEC RMS error.

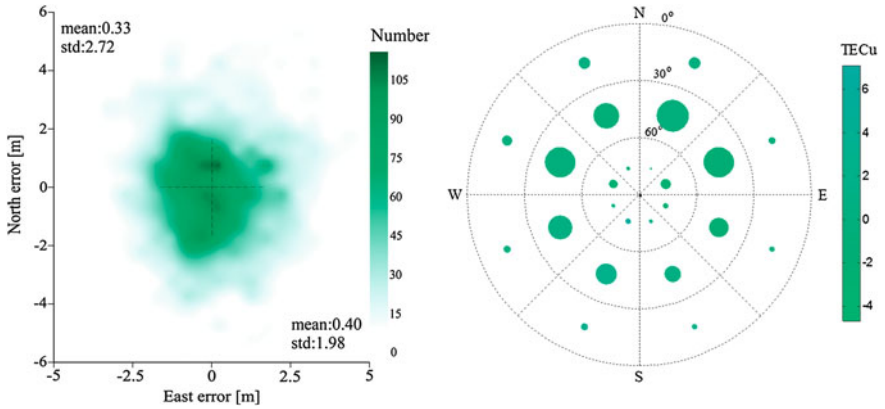


Fig. 39.13 Relation between the distributions of horizontal errors (*left plot*) and sTEC residual errors (*right plot*) for CODESH model at WUHN

However, there may appear exceptions sometimes. Figure 39.14 shows the relation between relative vTEC RMS errors and positioning 3D RMS in 2002 at all selected sites. It can be seen that the relation between 3D positioning RMS and relative vTEC error is obvious at the station of WUHN, LHAS and TWTF, whereas at the station of SHAO and KUNM, positioning errors do not follow the ionospheric correction accuracy trends.

We referred to the following interpretation to this phenomenon [5]. The positioning accuracy of single-frequency users depends mainly on whether the ionospheric biases of the corresponding ionospheric corrections are consistent for satellites in view. A constant bias would only affect the receiver clock error as a consequence of the column of ones in design matrix.

Based on ionospheric sTEC residual errors, which are calculated from the modeled and measured sTEC values, we defined sTEC bias as the difference between the maximal and minimal sTEC residual errors for satellites in view at the

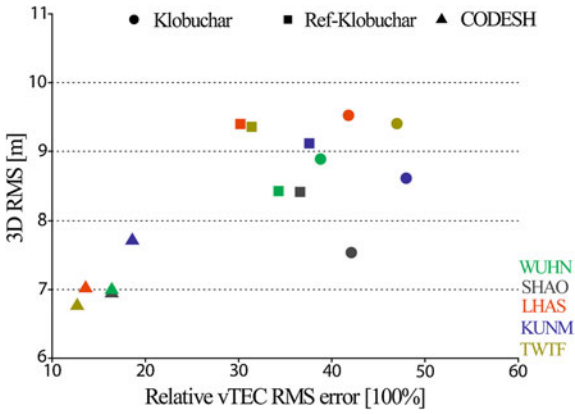


Fig. 39.14 The relation between relative vTEC RMS errors and positioning 3D RMS for the selected sites in 2002 (*circles* for Klobuchar, *squares* for Ref-Klobuchar, *triangles* for CODESH and *different colors* for the test sites)

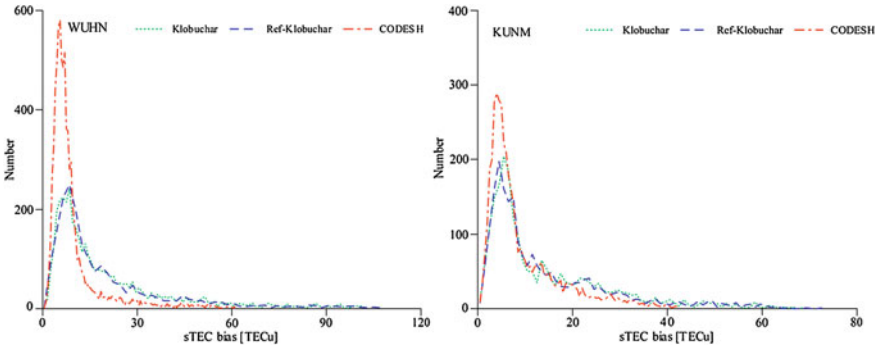


Fig. 39.15 Distributions of sTEC biases with the ionospheric corrections at the station of WUHN (*left plot*) and KUNM (*right plot*) in 2002

individual epoch. Figure 39.15 shows the distribution of sTEC biases of the corresponding ionospheric corrections at the station of WUHN and KUNM. In the left plot, it is stated that 95 % of the sTEC biases are less than 63 and 62 TECu for Klobuchar and Ref-Klobuchar algorithms respectively. As for CODESH model, it significantly decreases sTEC bias to 30TECu (in 95 % accuracy), which is much smaller with respect to the other models. Moreover, CODESH model also presents a better shape, so we can foresee a better positioning performance with this algorithm. In the right plot, although CODESH stands out, it is still hard to speculate which will have a better positioning result when using Klobuchar or Ref-Klobuchar model. Hence, the relation between ionospheric vTEC performances and the corresponding positioning errors should be further analyzed.

39.5 Conclusions

Focusing on the mid-low latitudes in China, the performances of different ionospheric corrections in terms of TEC correction accuracy and SPSE positioning are presented during periods of high (2002) and low (2009) solar activity. Better ionospheric TEC correction performances are obtained in 2002 than in 2009 (58, 64, and 84 % versus 38, 51 and 83 % for Klobuchar, Ref-Klobuchar and CODESH model respectively) as a result of the small mean measured vTEC value (6.6 TECu) during solar minimum. Hence, the PTRC (relative error) and TEC RMS (absolute error) are suggested to be employed together for evaluating the accuracy of ionospheric corrections in practice. As for positioning performances, significant vertical improvement is found with ionospheric corrections compared with the uncorrected case for both of the year 2002 and 2009 (increasing 24–57 % and 55–68 % respectively), while horizontal accuracy improves much less with respected to the vertical one (increasing 4–29 % in 2009). In addition, the horizontal error might become even large with ionospheric models (Klobuchar and

Ref-Klobuchar algorithms increase the horizontal error by about 27 % in 2002, performing even worse than the uncorrected case).

The relationship between ionospheric TEC performance and positioning error is discussed. It is stated that the correlation between them at some sites is obvious, implying that positioning accuracy is related to TEC modeling to some degree. The impacts of ionospheric corrections on GNSS positioning accuracy in particular for horizontal accuracy are also analyzed. The distribution of the uncorrected horizontal errors is characterized by a large dispersion along the north direction than along the east one at the mid-low latitudes in China, indicating the presence of north–south TEC gradients and a key role of north error in horizontal accuracy. Focusing on the station of WUHN, the performances of Klobuchar and CODESH algorithms on correcting the north–south TEC gradients are accessed, from which we highlight the impact of TEC horizontal gradients on north component.

Although ionospheric corrections can present good TEC performances and provide relatively high positioning accuracy with respect to the uncorrected case, these algorithms improve more the vertical accuracy than the horizontal one due to the presence of horizontal TEC gradients. Hence the impact of TEC gradients on horizontal positioning should be further investigated, which will benefit the establishment and assessment of the future COMPASS global broadcast ionospheric model.

Acknowledgments This work is supported by National 973 (No. 2012CB825604) and 863 programs (No. 2012AA121803), China Natural Science Funds (No. 41231064, 41104012, 41021003), the CAS/SAFEA International Partnership Program for Creative Research Teams (KZZD-EW-TZ-05) and the Knowledge Innovation Project of the CAS (KGCXZ-EW-407-2). The authors would like to thank the IGS Global Data Center CDDIS (Crustal Dynamics Data Information System, Greenbelt, MD, USA) for granting access to dual-frequency GPS observation data used in this study. We also thank the Center for Orbit Determination in Europe (CODE, University of Berne, Switzerland) for making the daily ionospheric coefficients and monthly DCB values available online.

References

1. Yuan YB, Huo XL, Ou JK (2007) Models and methods for precise determination of ionospheric delays using GPS. Prog Nat Sci 17(2):187–196
2. Klobuchar JA (1987) Ionospheric time-delay algorithm for single-frequency GPS users. IEEE Trans Aero Electron Syst AES 23(3):324–331
3. Arbesser-Rastburg B (2006) The galileo single frequency ionospheric correction algorithm. In: The 3rd European space weather week, Brussels, Belgium
4. Bidaine B, Warnant R (2011) Ionosphere modelling for Galileo single frequency users: illustration of the combination of the NeQuick model and GNSS data ingestion. Adv Space Res 47(2):312–322
5. Fees WA, Stephens SG (1987) Evaluation of GPS Ionospheric time-delay model. IEEE Trans Aero Electron Syst AES 23(3):332–338
6. Somieski A, Burgi C, Favey E (2007) Evaluation and comparison of different methods of ionospheric delay mitigation for future galileo mass market receivers. In: ION GNSS 2007, 25–28 Sep, Fort Worth, Texas, USA, pp 2854–2860

7. Jakowski M, Hoque MM (2012) Ionospheric range error correction models. In: ICL-GNSS 2012, 25-27 June, Starnberg, Germany, pp 1–6
8. Yang Z, Song SL, Xue JC et al (2012) Accuracy assessment of Klobuhcar model and NeQuick model in China (in Chinese). Geom Inf Sci Wuhan Univ 37(6):704–708
9. Mohino E (2008) Understanding the role of the ionospheric delay in single-point single-epoch GPS coordinates. J Geodesy 82(1):31–45
10. Bidaine B, Lonchay M, Warnant R (2012) Galileo single frequency ionospheric correction: performances in terms of position. GPS Solution 2012:1–11
11. Chen JP, Wu B, Hu XG, et al. (2012) Compass/beidou: system status and initial service. In: IGS workshop on GNSS biases 2012, 18–19 Jan, University of Bern, Switzerland
12. Andrei C-O, Chen R, Kuusniemi H, et al, (2009) Ionosphere effect mitigation for single-frequency precise point positioning. In: ION GNSS 2009, 22–25 Sept, Savannah, Georgia, USA, pp 2508–2517
13. Schaer S (1999) Mapping and predicting the earth's ionosphere using the global positioning system. Doctor Dissertation. Switzerland: Astronomical Institute, University of Berne, pp 64–65
14. Li ZS, Yuan YB, Li H, et al. (2012) Two-step method for the determination of the differential code biases of COMPASS satellites. J Geodesy 2012:1–18
15. Macalalad PE, Tsai L-C, Wu J, et al. (2012) Application of the TaiWan ionospheric model to single-frequency ionospheric delay corrections for GPS positioning. GPS Solution 2012:1–10

Chapter 40
GWMT Global Atmospheric Weighted Mean Temperature Models: Development and Refinement

Changyong He, Yibin Yao, Dong Zhao, Ke Li and Chuang Qian

Abstract The atmospheric weighted mean temperature T_m plays a pivotal role in remote sensing water vapor with GNSS technique. However, real-time T_m cannot be obtained at given place due to limitation of modern technology. In order to remove this barrier, a global and season-specific T_m model (GWMT-I) based on spherical harmonics analysis was built using land radiosonde data. Combining Global Pressure and Temperature (GPT) model with the Bevis formula to provide virtual T_m observation at sea, a modified T_m model (GWMT-II) was presented aimed to improve the accuracy of GWMT-I in the ocean areas. Thereafter, the third generation T_m (GWMT-III) model, derived from GGOS Atmosphere surface global grid T_m data, came into existence. But on the one hand, these three models fail to accurately describe vertical T_m variation. On the other hand, they are expanded into spherical harmonics up to degree and order nine which may impose restriction to their precision. This paper reconstructs the latest model (GWMT-IV) with 4D global NCEP/DOE 2 Reanalysis data from 2005 to 2009. Comparison of result between 2010 NCEP2 data and the GWMT-IV model shows that it achieves high accuracy all over the world and its modeling method better conforms to reality.

Keywords GNSS meteorology · Weighted mean temperature · GWMT models · GWMT-IV model

C. He (✉) · Y. Yao · D. Zhao
School of Geodesy and Geomatics, Wuhan University, Wuhan, China
e-mail: ChangY.He@gmail.com58707019@qq.com

K. Li
College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China

C. Qian
GNSS Research Centre, Wuhan University, Wuhan, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_40,

40.1 Introduction

Water vapor is a vital component of Earth atmosphere playing a crucial role in globally atmospheric radiation, hydrological cycle and energy equilibrium. Remotely measuring the Precipitable Water Vapor (PWV, the height of an equivalent column of liquid water) [1] is a staple part for forecasting short-term rainstorm and Meiyu season [2], for improving numerical weather prediction [3, 4], for studying tempo-spatial distribution and diurnal variations of PWV [5], for monitoring severe weather events including thunderstorms, hail stones, strong wind and hurricanes [6] and for studying global climate change [7].

The tropospheric zenith total delay (ZTD) is firstly estimated in remote sensing PWV with GPS technology. The zenith 'wet' delay (ZWD), which is mainly contributed to the water vapor in the atmosphere, can be separated from ZTD by Saastamoinen model with measured air pressure [8]. The relationship between ZWD and PWV is presented by Askne and Nornius in 1987 which led to the ideas of detecting atmosphere with ground-based GPS [9, 10]. The relationship between PWV and ZWD is expressed as follows:

$$PWV = \Pi \cdot ZWD = \Pi \cdot (ZTD - ZHD) \tag{40.1}$$

$$\Pi^{-1} = 10^{-6} \cdot \rho_w R_v \left[k_3/T_m + k_2'\right] \tag{40.2}$$

$$T_m = \frac{\int (e/T) Z_w^{-1} dz}{\int (e/T^2) Z_w^{-1} dz} \tag{40.3}$$

$$Z_w^{-1} = 1 + 1650 \frac{e}{T^3}\left(1 - 0.01317t + 1.75 \times 10^{-4} t^2 + 1.44 \times 10^{-6} t^3\right) \tag{40.4}$$

where $\prod$ is the conversion factor for water vapor, ρ_w is the density of liquid water and water vapor, R_v is the gas constant for water vapor, k_2' and k_3 are atmospheric refractivity constants [11], e is the water vapor pressure of a unit in hPa, T and T_m are the absolute temperature and weighted mean temperature in units of Kelvin. Z_w^{-1} is the inverse compressibility factor which is correction for non-ideal gas [12]. The error of $\prod$ is the main resource in sensing PWV with GPS when we neglect the error of ZWD [13].

In practical application, it is hard to accurately estimate the real-time T_m by the Eq. (40.3) with the given data of atmospheric temperature and water vapor. Currently, there are three reliable means to compute T_m [14]:

1. integral method based on radiosonde and numerical weather prediction data; The limit tempo-spatial resolution of these data produces the barrier for its application in GPS meteorology. Due to low temporal resolution or poor coverage of these data, this method is rarely used in real-time PWV estimation.
2. regression method; The most widely used method is to derive a statistical relationship between T_m and T_s, the surface temperature. Bevis et al. [15] at the University of Hawaii firstly gave a linear regression formula for estimating

mid-latitude T_m using radiosonde data from 13 stations in North America, with RMS (Root-Mean-Square Error) about 4.74 K (the Kelvin scale). Mendes et al. [16, 17] proposed linear and polynomial relationship between T_s and T_m according to the radiosonde data of 50 sites covering the latitude range of 62°S to 83°N and a height range of 0–2.2 km. The analysis of radiosonde soundings at 53 globally distributed stations given by Ross and Rosenfeld, showed that the correlation between T_s and T_m varies with geographical position and season, and becomes weaker in the tropics [18]. In China, Li [19] established $T_s - T_m$ regression formula for Chengdu. Wang et al. [20] noted that T_s and T_m share very significant relationship. Wang et al. [21] revealed seasonal variations of regression coefficients in Bevis model using Hong Kong Park sounding station data from 2003 to 2009.

3. non-meteorological empirical models used as input the site latitude, longitude and the time; Emardson et al. [22] developed an annual model, which is applicable to most parts of the Europe, to obtain conversion factor $\prod$ (defined as function of T_m) using European Radiosonde data. The validation of Mendes et al. [16, 17] showed that the Emardson model performs poorly and the use of T_s is necessary in obtaining a higher accuracy in T_m determination. Yao et al. [23] established a globally applicable, season-specific GWMT-I model for estimating T_m using land radiosonde data with a global average RMS of less than 6 K. In order to address its anomalies in the oceanic areas, the GWMT-II was developed. The model coefficients were recalculated and GWMT-III was built using 2.5° × 2.5° global T_m grid data provided by GGOS Atmosphere (Global Geodetic Observing System). The global average RMS of GWMT-III model is ∼3.4 K. But these GWMT models were built using surface T_m data and regarding global lapse rate of T_m as constant. Moreover, the GWMT models were based on degree and order nine spherical harmonics expansion which may severely limit their accuracy.

In this paper we develop GWMT-IV model, which is aimed to provide better estimation of T_m over the world, using NCEP-DOE 2 (The National Centers for Environmental Prediction) data (referred to as NCEP2). We firstly introduce some details of existing GWMT models and the modeling methodology of the GWMT-IV model with global NCEP2 data from 2005 to 2009. A validation of their performances with some reliable references derived from NCEP 2 and GGOS Atmosphere data in 2010 is presented in the end of this paper.

40.2 Existing Empirical Models Description

Existing T_m models include Emardson model, the GWMT-I, the GWMT-II and the GWMT-III. One of these models uses the site latitude and the time of the year as input, other three models use latitude, longitude and the time of the year.

Table 40.1 Values for Coefficients used in Eq. (40.5)

Coefficients	Value
a_0	5.861
a_1	0.011
a_2	0.054
a_3	0.138

40.2.1 Emardson Model

Emardson et al. [22] generated an annual model to obtain $\prod$ using European Radiosonde data from Scandinavia.

$$\Pi^{-1} = a_0 + a_1\theta + a_2 \cdot \sin\left(2\pi\frac{doy}{365}\right) + a_3 \cdot \cos\left(2\pi\frac{doy}{365}\right) \tag{40.5}$$

where θ is the latitude in degrees and *doy* is the day of the year. The coefficients for this model are presented in Table 40.1. Substituting Eq. (40.5) into Eq. (40.2), we can compute T_m. The Emardson model is still chosen as an empirical model for estimating T_m, although it is only valid for most European continent.

40.2.2 GWMT-I, GWMT-II and GWMT-III Models

On the foundation of the GPT [24] model and Bevis T_s– T_m relationship, GWMT-I model is based on spherical harmonics analysis using geodetic coordinates of site as input parameters:

$$T_m(\varphi, \lambda) = \alpha_1 + \alpha_2 h + \alpha_3 \cos(\frac{doy - 28}{365.25}2\pi) \tag{40.6}$$

where φ, λ and h are the latitude, longitude and height of stations, respectively. α_i (i = 1 or 3) are expended into spherical harmonics up to degree nine and order nine, but the globally mean lapse rate of T_m (α_2) is -0.0041 K/*km*:

$$\alpha_i = \sum_{n=0}^{9}\sum_{m=0}^{n} P_{nm}(\sin\varphi) \cdot \left[A_{nm}^i \cos(m\lambda) + B_{nm}^i \sin(m\lambda)\right] \tag{40.7}$$

where P_{nm} is Legendre polynomials. A_{nm}^i and B_{nm}^i, the coefficients for degree n and order m in Eq. (40.7), have been estimated using least squares method.

In order to address its anomalies in the oceanic areas, the GPT (Global Pressure and Temperature) model was combined with the Bevis T_m–T_s formula to provide virtual T_m observations at the sea and the GWMT-II was developed. For the sake of further improving accuracy, the GWMT model coefficients were re-determined and GWMT-III was constructed using global T_m grid data provided by GGOS Atmosphere.

40.3 Determination of GWMT-IV

The determination of GWMT-IV is based on five-year (January 2005 to December 2009) NCEP2 reanalysis data. NCEP2 data have a vertical resolution of 17 pressure levels (i.e., 1,000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa) and a horizontal resolution of 2.5° × 2.5° at 6 h temporal resolution. The pressure-level data used for T_m modeling include temperature, geopotential height, pressure and humidity over the period of 2005–2009. T_m can be computed from these layer meteorological profiles via numeric integrals in Eq. (40.3).

It is unreasonable to assume that the lapse rate of T_m is invariable worldwide. Therefore, the global lapse rate of T_m is parameterized dependent on geodetic coordinates and time. The GWMT-IV model developed in this paper is applicable to the heights between 0 and 7 km. Assuming the T_m at the given grid point and time decreases linearly with height, we can get

$$T_m(\varphi, \lambda, h) = \beta_m(\varphi, \lambda) \cdot h + T_m^0(\varphi, \lambda) \tag{40.8}$$

where $\beta_m(\varphi, \lambda)$ is the lapse rate of T_m at the latitude of φ and the longitude of λ. $T_m^0(\varphi, \lambda)$, $T_m(\varphi, \lambda, h)$ are the T_m at the mean sea level and the orthometric height of h (km). After analyzing T_m^0 and β_m time series, GWMT-IV uses similar method to treat them. As an example see Eq. (40.9) for T_m^0.

$$T_m^0(\varphi, \lambda) = mean + amp_1 \cdot \cos(2\pi \frac{doy - D_1}{365.25}) + amp_2 \cdot \cos(4\pi \frac{doy - D_2}{365.25}) \tag{40.9}$$

In Eq. (40.9), $mean$, amp_1 and amp_2 are the annual mean value, annual and semiannual variations of T_m at mean sea level. Similar to the GPT model given by Boehm et al. [24], the phases of annual variation in former GWMT models is fixed to January 28 which may severely limit their accuracy. Aiming to improve the accuracy of GWMT-IV model, we regard the phases of annual variation (D_1) and semiannual variation (D_2) of T_m at mean sea level as parameters dependent on location.

In the next step, $mean$, amp_1 and amp_2 are expanded into spherical harmonics up to degree and order N_{max}. For instance, $mean$ is described in Eq. (40.10). Unlike Eq. (40.7), $\tilde{P}_{nm}$in Eq. (40.10) is the fully normalized associated Legendre polynomial [25] in order to compute correspondent coefficients $\tilde{A}_{nm}$ and $\tilde{B}_{nm}$ up to degree and order more than 30 using least square method.

$$mean = \sum_{n=0}^{N\max} \sum_{m=0}^{n} \tilde{P}_{nm}(\sin\varphi) \cdot \left[\tilde{A}_{nm}\cos(m\lambda) + \tilde{B}_{nm}\sin(m\lambda)\right] \tag{40.10}$$

Figure 40.1 presents the fitting RMS and coefficient of determination for T_m^0 (Fig. 40.1a) and β_m (Fig. 40.1b) as a function of N_{max}. As can be seen, the fitting error decreases with increasing N_{max}. On the contrary, the coefficient of

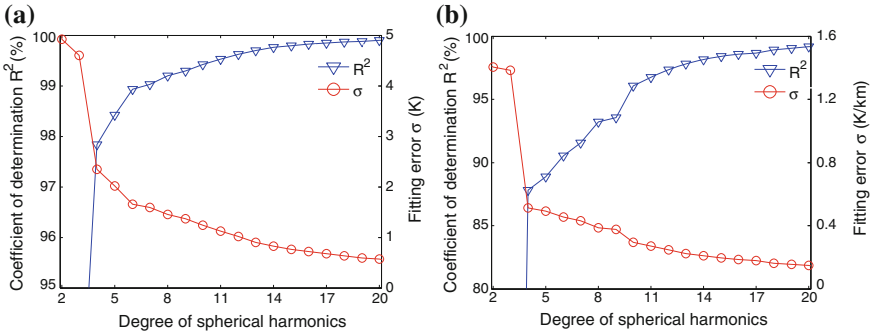


Fig. 40.1 The coefficient of determination and fitting error for (**a**) T_m^0 and (**b**) β_m

determination rapidly increases more than 95 %. Consequently, the N_{max} is set to 15 taking the complexity of the model into account.

40.4 Validations of T_m Models

To verify the performance of GWMT-IV and the former GWMT models, an validation process is conducted. In this study, NCEP2- and GGOS-derived T_m were chosen as references, and bias and RMS are used to evaluate the accuracy and stability of the models. The formulas are given as below.

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(T_m^i - \tilde{T}_m^i) \tag{40.11}$$

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(T_m^i - \tilde{T}_m^i)^2} \tag{40.12}$$

where T_m, $\tilde{T}_m$ and N are the T_m prediction derived from empirical models, the reference T_m value and the total number of measurements during the experimental period of time. Because of the highest accuracy of GWMT-III among three former GWMT models, only the Emardson model, the GWMT-III and the improved GWMT-IV are selected in the next comparisons.

40.4.1 Comparison with GGOS Data

The GWMT-III, GWMT-IV and the Emardson are selected in this comparison. The GWMT-III was constructed based on the surface T_m data provided by GGOS,

Table 40.2 Statistics of global bias and RMS for three T_m empirical models compared with GGOS data (Unit: K)

Models	Bias	RMS
Emardson	29.1 [−10.4, 128.2]	32.6 [2.5,128.9]
GWMT-III	−0.1 [−4.7, 4.9]	3.2 [1.1, 6.8]
GWMT-IV	3.7 [−4.1,11.2]	5.4 [1.5, 11.9]

Values in square brackets show the minima and maxima of both bias and RMS, the same as following tables

hence the surface T_m data are used to examine the accuracy of these three models as well. Compared with daily GGOS T_m data (namely at 0, 6, 12, 18 UTC), Table 40.2 shows the statistics of global bias and RMS for these three models, and Fig. 40.2 illustrates the distribution of comparison results in 2010.

Compared with the statistics of GGOS data, Table 40.1 indicates that the result of global bias and RMS for Emardson model for around 30 K is far from satisfactory. Especially, in Antarctic continent with the south of 60° S, the RMS can climbe up to 120 K (see Fig. 40.2a and b). It also can be seen from the table that the GWMT-III is the closest to GGOS T_m data among these models with worldwide maximal RMS for less than 7 K and global average RMS for 3.2 K which are in accordance with the result described in the introduction. The global bias of the GWMT-III (GWMT-IV) is −0.1 K (3.7 K). In Fig. 40.2e and f, the largest biases are mainly concentrated on the latitude belt between 30° and 60° in northern and southern hemispheres, especially in the sea of the southern hemisphere where both bias and RMS values jump to over 10 K.

There are some reasons described as below to suppose to explain this phenomenon could be. On the one hand, the GGOS data are recorded in different pressure layers and therefore interpolation is required to get the meteorological parameters at the specific height. On the other hand, the air pressure at the bottom of reanalysis data is usually 1,000 hPa (the air pressure over the sea is around 1,010 hPa). Thus extrapolation is demanded to get meteorological parameters near the mean sea level.

40.4.2 Comparison with NCEP2 Data

The global NCEP2 data of 2010 in 925 hPa (∼500 m) and 600 hPa (∼4.5 km) layers, which is not introduced into GWMT-IV modeling, is used to justify the reliability of GWMT-IV. The Emardson model will not be verified in the following comparisons due to its lowest accuracy among three models. Tables 40.3 and 40.4 give the statistics information of bias and RMS compared with NCEP2 data in different altitude for GWMT-III and GWMT-IV. Figures 40.3 and 40.4 show their global color coded maps of bias and RMS.

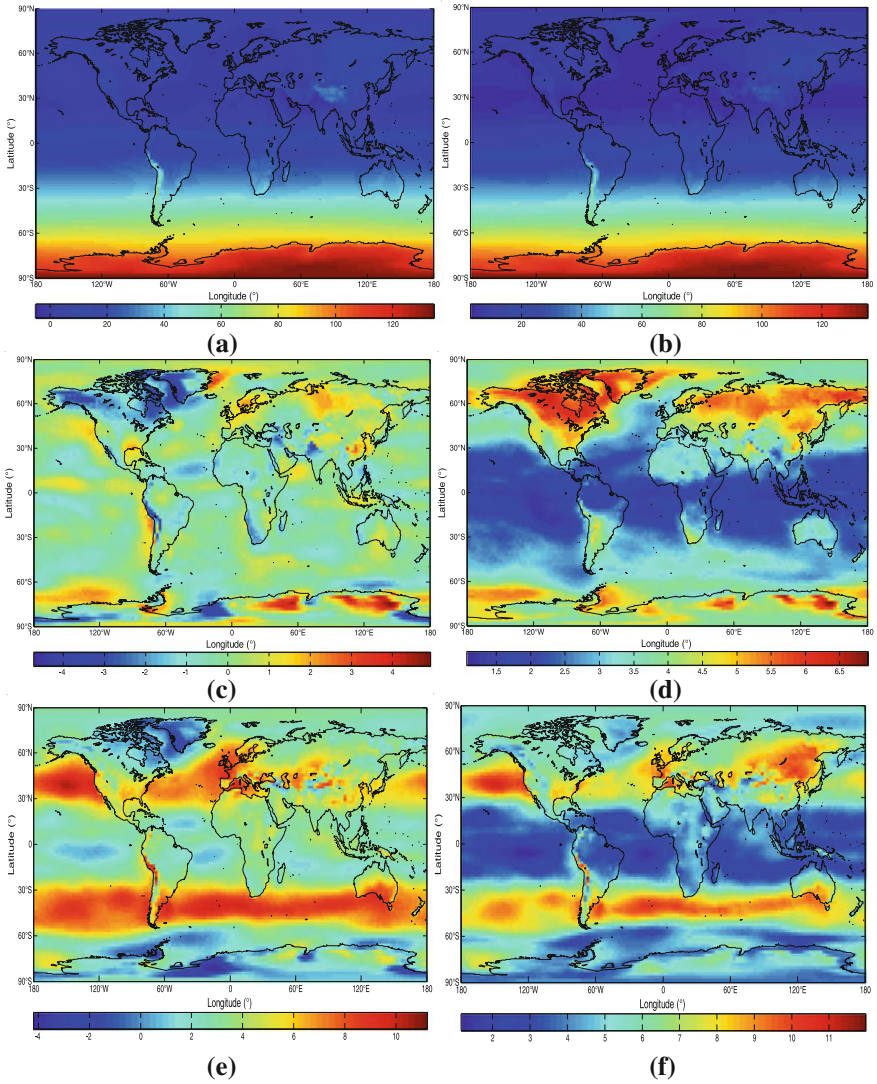


Fig. 40.2 Comparison of global T_m between GGOS and three models in 2010. **a** Bias of Emardson (K). **b** RMS of Emardson (K). **c** Bias of GWMT-III (K). **d** RMS of GWMT-III (K). **e** Bias of GWMT-IV (K). **f** RMS of GWMT-IV (K)

Table 40.3 indicates the comparison result with NCEP2 data in 925 hPa layer. It can be found that GWMT-III is worese fit to NCEP2 reanalysis data than GWMT-IV, whose global average bias (RMS) is −0.03 K (4.4 K). Figure 40.3 illustrates the distribution of bias and RMS for these two models, and it can be seen that the improved T_m model is significantly better than the GWMT-III except for the Siberia and Far East Area.

Table 40.3 Statistics of global bias and RMS compared with NCEP2 data in 925 hPa layer (Unit: K)

Models	Bias	RMS
GWMT-III	−3.6 [−9.2, 10.8]	6.1 [2.0, 16.4]
GWMT-IV	−0.02 [−6.2, 7.4]	4.1 [0.8, 12.9]

Table 40.4 Statistics of global bias and RMS compared with NCEP2 data in 600 hPa layer (Unit: K)

Models	Bias	RMS
GWMT-III	−2.9 [−15.1, 8.0]	8.1 [3.4, 15.3]
GWMT-IV	−0.3 [−4.3, 7.4]	5.6 [0.9, 13.6]

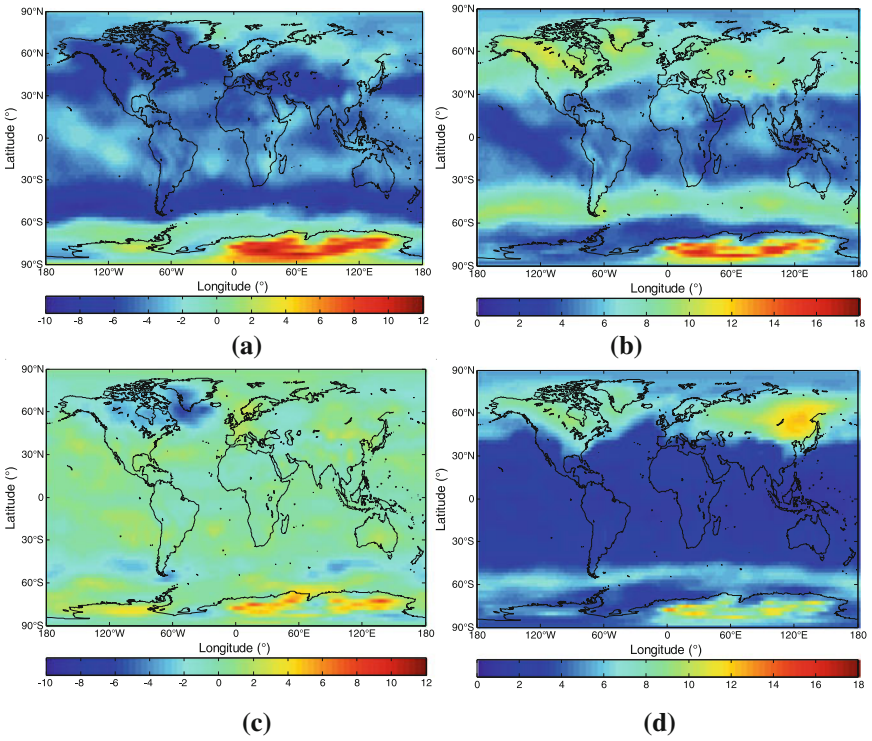


Fig. 40.3 Comparison of global T_m between NCEP2 and two models in 925 hPa layer in 2010. **a** Bias of GWMT-III (K). **b** RMS of GWMT-III (K). **c** Bias of GWMT-IV (K). **d** RMS of GWMT-IV (K)

In addition, we validate these models in 600 hPa layer as well (see Table 40.4 and Fig. 40.4). The accuracy of GWMT-IV is almost higher than that of GWMT-III over the world. Compared with the result in 925 hPa layer, the accuracy of two

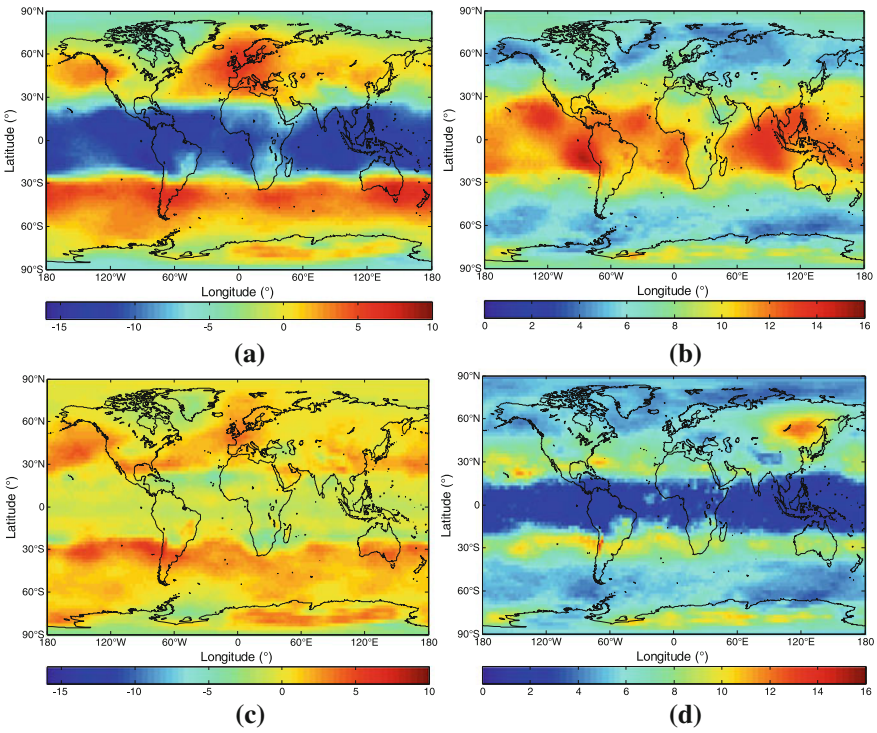


Fig. 40.4 Comparison of global T_m between NCEP2 and two models in 600 hPa layer in 2010. **a** Bias of GWMT-III (K). **b** RMS of GWMT-III (K). **c** Bias of GWMT-IV (K). **d** RMS of GWMT-IV (K)

models generally decline along with altitude. For example, accuracy of the GWMT-III falls fastest in tropics (bias: ~ -12 K, RMS: ~ 13 K) which is mainly caused by the constant lapse rate of T_m around the world. Figure 40.5 presents the global distribution of lapse rate of T_m for GWMT-IV. It can be concluded that absolute value of lapse rate of T_m is large in tropical regions and is small in temperate and polar region. Conversely, GWMT-IV achieves a high accuracy between the thirty parallels (RMS < 4 K), which effectively justifies that the parameterization of lapse rate of T_m can evidently improve the accuracy of T_m models.

The comparisons of GWMT-IV with GGOS and NCEP2 data reveal the unignorable difference in T_m between GGOS and NCEP2, which means more reliable T_m data resources are required to develop more accurate GWMT models. In brief, the global average RMS of non-meteorological empirical T_m models is approximately 3–6 K.

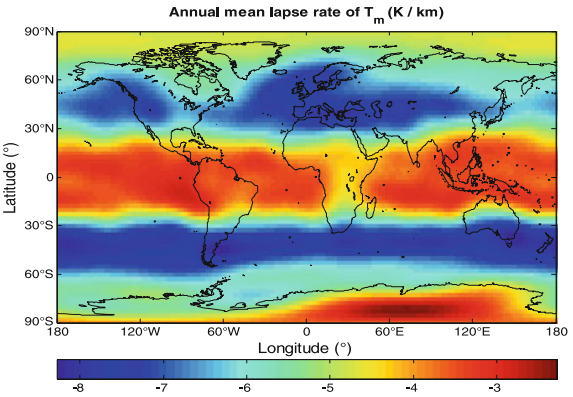


Fig. 40.5 Global distribution of annual mean lapse rate of T_m

40.4.3 Comparison with Bevis T_s-T_m Relationship

The comparisons above present the advantages of GWMT-IV among these three models. In order to comprehensively test the GWMT-IV, GWMT-III, GWMT-IV are compared with the Bevis T_m-T_s relationship. The Bevis T_m–T_s relationship ($T_m = a + bT_s$) on each grid is determined from the analysis of daily NCEP2 profiles in 2010. Similarly, the result is illustrated as a function of latitude in 925 hPa and 600 hPa layer (see Fig. 40.6).

Figure 40.6 indicates that T_m estimations from GWMT-IV are closer to that of Bevis T_s-T_m relationship. Generally, the difference of result between GWMT-III and GWMT-IV is obviously large in tropics and relatively small in cold and temperate zones. In 925 hPa layer (see Fig. 40.6a), the RMS of GWMT-IV is almost 2 K larger than that of GWMT-III in latitudes of 60°N to 60°S. Moreover, in the Fig. 40.6b, the maxima difference of RMS for these two models is over 8 K near the equator. This result conforms to the former conclusion that the accuracy of GWMT models is considerably decrease with the increasing heights.

40.4.4 Validation of PWV

We so develop the GWMT-IV model as to improve the accuracy of PWV estimation without measured meteorological parameters in GNSS meteorology. A comparison of PWV is applied to validate its real application. Figure 40.7 shows the result of global grid RMS.

This comparison is conducted described as follows. PWV estimation is calculated by ZWD deriving from integral of NCEP2 data and by T_m from the GWMT-IV model. The true PWV values are directly integrated from 1,000 hPa to 10 hPa layers by NCEP2 data of 2010. The global distribution of RMS in Fig. 40.7 shows that the differences between PWV estimation and true PWV value are basically smaller than 1.5 mm, less than 0.6 mm at high latitudes and $\sim$1 mm in

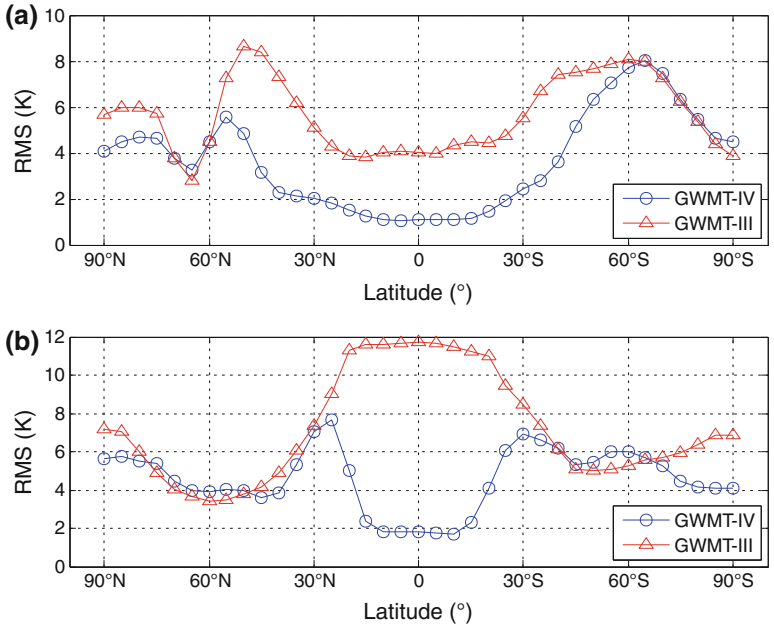


Fig. 40.6 Comparison of T_m between Bevis T_m–T_s relationship and two models in (**a**) 925 hPa and (**b**) 600 hPa layers using NCEP2 data in the year of 2010

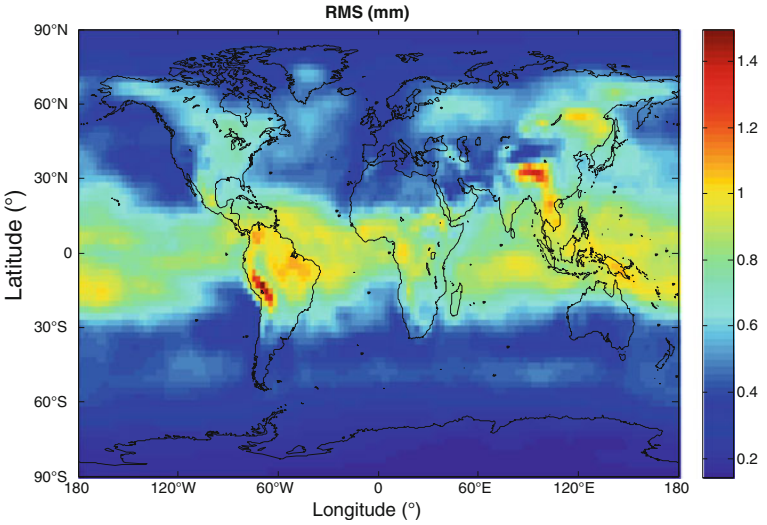


Fig. 40.7 Comparison of PWV in 1,000 hPa layer using NCEP2 data in the year of 2010

tropical areas. Notwithstanding, the RMS near the Himalaya Mountains and the Andes can reach up to 1.4 mm which means the accuracy of PWV for empirical models is relatively worse in these regions.

40.5 Conclusions

Considering the annual and semiannual variations of T_m and its lapse rate, GWMT-IV model is built in this paper based on spherical harmonics up to degree and order fifteen, using global T_m acquiring from the NCEP2 data, to improve the accuracy of GWMT models. Some conclusions can be drawn as below.

It is obvious that the accuracy of the GWMT-III and the GWMT-IV is generally better than that of the Emardson model. And the GWMT-III fit a little better to GGOS T_m data than the GWMT-IV. However, in the comparisons with NCEP2, the GWMT-IV model is far superior to the GWMT-III, particularly in mountain areas. Additionally, the T_m estimation of GWMT-IV is closer to the Bevis formula than that of GWMT-III comparing with NCEP2 data (especially in the tropics). Finally, the validation of PWV shows that the RMS error of PWV result from GWMT-IV is less than 2 mm.

The model coefficients in GWMT-IV had better be determined using NCEP2 data of the latest five years. The other data (e.g. radiosonde and COSMIC data) may be required to further assess these GWMT models. It is worth noting that the GWMT-IV is better match up to the reality in theory and application, and this modeling method has its certain guiding significance as well.

Acknowledgments We would like to thank the NOAA NCEP/NCAR for providing global NCEP-DOE Reanalysis 2 data from 2005 to 2010 and the GGOS Atmosphere for providing access to surface T_m data in 2010. We are also very grateful to the reviewers for their valuable advices to improve the manuscript.

References

1. Duan J, Bevis M, Fang P, et al. (1996) GPS meteorology: direct estimation of the absolute value of precipitable water. J Appl Meteorol 35(6):830–838. doi: 10.1175/1520-0450(1996)035<0830:GMDEOT>2.0.CO;2
2. Song S, Zhu W, Din J, Cheng Z (2003) Coming process of meiyu deason in 2002 over Yangtz River delta monitored by Shanghai GPS network. Prog Astron 21(2):180–184
3. Kuo YH, Guo YR, Westwater ER (1993) Assimilation of precipitable water measurements into a mesoscale model. Mon Weather Rev 121:1215–1238
4. Gutman SI, Sahm S, Stewart J, Benjamin S, Smith T (2003) A new composite observing strategy for gps meteorology. In: 12th symposium on meteorological observations and instrumentation 2003, American Meteorology Society, Long Beach, California
5. Wu PM, Hamada JI, Mori S, Tauhid YI, Yamanaka MD, Kimura F (2003) Diurnal variation of precipitable water over a mountainous area of Sumatra Island. J Appl Meteorol 42:1107–1115. doi:10.1175/1520-0450(2003)042<1107:DVOPWO>2.0.CO;2

6. Choy S, Zhang K, Wang C-S, Li Y, Kuleshov Y (2011) Remote sensing of the earth's lower atmosphere during severe weather events using GPS technology: a study in Victoria, Australia. In: Proceedings of ION GNSS 2011, 20-23 Sept, Portland, Oregon, pp 559–571
7. Gradinarsky LP, Johansson JM, Bouma HR, Scherneck HG, Elgered G (2002) Climate monitoring using GPS. Phys Chem Earth 27:335–340. doi:10.1016/S1474-7065(02)00009-8
8. Saastamoinen J (1972) Contributions to the theory of atmospheric refraction. Part II Refraction Corrections Satellite Geodesy Bull Géodésique 107:13–34
9. Askne J, Nornius H (1987) Estimation of tropospheric delay for microwaves from surface weather data. Radio Sci 22(3):379–386
10. Davis JL, Herring TA, Shapiro II, Rogers AE, Elgered G (1985) Geodesy by radio interferometry: effects of atmospheric modeling errors on estimates of baseline length. Radio Sci 20:1593–1607
11. Bevis M, Businger S, Chiswell S (1994) GPS meteorology: mapping zenith wet delays on to precipitable water. J Appl Meteorol 33:379–386. doi:10.1175/1520-0450(1994)033<0379:GMMZWD>2.0.CO;2
12. Owens JC (1967) Optical refractive index of air: dependence on pressure, temperature and composition. Appl Opt 6:51–59
13. Cheng J (1998) On the error analysis for the remote sensing of atmospheric water vapor by ground based GPS. Acta Geodaetica et Cartographica Sinica 27(2):113–118
14. Schueler T, Posfay A, Hein GW, Biberger R (2001) A global analysis of the mean atmospheric temperature for GPS water vapor estimation. In: ION GPS 2001 14th international technical meeting of satellite division of the Institute of Navigation, Salt Lake City, Utah
15. Bevis M, Businger S, Herring TA, Rocken C, Anthes RA, Ware RH (1992) GPS meteorology: remote sensing of atmospheric water vapor using the global positioning system. J Geophys Res 97(D14):15787–15801
16. Mendes VB (1999) Modeling the neutral-atmosphere propagation delay in radiometric space techniques. PhD dissertation, Department of Geodesy and Geomatics Engineering Technical Report No. 199, University of New Brunswick, Fredericton, New Brunswick, Canada 1999
17. Mendes VB, Prates G, Santoa L, Langley RB (2000) An evaluation of the accuracy of models for the determination of the weighted mean temperature of the atmosphere. In: Proceedings of ION 2000, National Technical Meeting, Anaheim, California, USA, pp 433–438
18. Ross RJ, Rosenfeld S (1997) Estimating mean weighted temperature of the atmosphere for global positioning system applications. J Geophys Res 102(D18):21719–21730. doi: 10.1029/97JD01808
19. Li G (2007) Detecting precipitable water vapor by ground-based GPS and its applications in meteorology. PhD dissertation, Southwest Jiaotong University 2007
20. Wang Y, Yang B, Liu Y, Hu X, Liu L, Xu H (2010) The study of the model about mean vapor pressure-weighted temperature of the atmosphere based on radiosonde. Sci Survey Mapp 35(2):112–113
21. Wang X, Dai Z, Cao Y, Song L (2011) Weighted mean temperature T_m statistical analysis in ground-based GPS in China. Geomatics Inform Sci Wuhan Univ 36(4):412–416
22. Emardson TR, Elgered G, Johansson JM (1998) Three months of continuous monitoring of atmospheric water vapor with a network of global positioning system receivers. J Geophys Res 103:1807–1820. doi:10.1029/97JD03015
23. Yao Y, Zhu S, Yue S (2012) A globally applicable, season-specific model for estimating the weighted mean temperature of the atmosphere. J Geodesy 86(12):1125–1135. doi:10.1007/s00190-012-0568-1
24. Boehm J, Heinkelmann R, Schuh H (2007) Short note: a global model of pressure and temperature for geodetic applications. J Geodesy 81(10):679–683. doi:10.1007/s00190-007-0135-3
25. Abramowitz M, Stegun IA (1965) Handbook of mathematical functions. Dover Publications, NY

Chapter 41
The Mathematical Expectation of GDOP and its Application

Tao Han, Haitao Wu, Xiaochun Lu, Juan Du and Xiaozhen Zhang

Abstract Along with the development of Global Navigation Satellite System (GNSS) and the deepening of research on Interoperability, more navigation satellites in space are able to provide Position Velocity and Time (PVT) service for users. Although added visible satellite number leads a better Geometric Dilution of Precision (GDOP), there is not a conclusion about this GDOP improvement in quality. In this paper, we obtain another expression of GDOP using two variables, azimuth and elevation, instead of the used three direction cosines. Then, suppose the two variables as random variable, we obtain the expression of the mathematical expectation of GDOP, which means the average GDOP when the satellites randomly appear in the visible space. After that, considering the different block situation of user, the minimal elevation angle is imported as a constraint condition, and we obtain the expression of GDOP mathematical expectation with the minimal elevation angle limitation. Furthermore, a simulation process is adopted to verify the foregoing expressions. At last, an assessment model of GNSS constellation using the GDOP mathematical expectation is constructed and which is employed to evaluate the constellation of interoperability between different GNSSs.

T. Han (✉) · X. Lu · J. Du · X. Zhang
National Time Service Center, Chinese Academy of Sciences, Xi'an, People's Republic of China
e-mail: hantao@ntsc.ac.cn

T. Han
Key Laboratory of Precision Navigation and Timing Technology, Xi'an, People's Republic of China

T. Han
Graduate University of the Chinese Academy of Sciences, Beijing, People's Republic of China

H. Wu
Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing, People's Republic of China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_41,

Keywords GDOP · Mathematical expectation · Uniform distribution · Assessment model

41.1 Introduction

Interoperability and compatibility have been hot-spot topic ever since the Galileo was planned, and recently, the importance of these concepts is even increasing more and more. In 2008, at the third International Committee on Global Navigation Satellite System (ICG), global and regional system providers agreed that at a minimum, all global navigation satellite systems (GNSS) signals and services must be compatible. To the maximum extent possible, open signals and services should also be interoperable, in order to maximize benefit to all GNSS users. The latest definitions of interoperability and compatibility are as below [1]:

> Interoperability refers to the ability of global and regional navigation satellite systems and augmentations and the services they provide to be used together to provide better capabilities at the user level than would be achieved by relying solely on the open signals of one system.

And:

> Compatibility refers to the ability of global and regional navigation satellite systems and augmentations to be used separately or together without causing unacceptable interference and/or other harm to an individual system and/or service.

Nowadays, multi-GNSS age is coming and it will provide more visible satellites for users. This increasing visible satellites will bring both benefits and harmful. In one hand, adding visible satellites will provide more pseudorange measurements for coordinates solving process; this will ameliorate the Geometric Dilution of Precision (GDOP) of navigation solution, and lead a more accurate solution. In another hand, adding visible satellites provide more service signals that will largen interference between the GNSS systems, furthermore, increase receiver noise floor. The interference is not only caused by intra-system signals, but also caused by inter-system signals [2].

That is to say, adding visible satellite number bring both benefits and harmful on solution accuracy. In order to measure the impact of increasing visible satellite number on solution accuracy, the GDOP average is required. The GDOP average means that when the visible satellites are randomly appears in the space, the GDOP value can be obtained for position solution in average. To calculate the GDOP average, the location of each visible satellite can be seen as a random variable. Thus, the GDOP average can also called “Mathematical Expectation of GDOP Value”. The purpose of this paper is to calculate the expression of this “mathematical expectation”.

41.2 Mathematical Expectation of GDOP Value

Navigation satellite systems utilize pseudorange measurement for navigation solution. This method assumes receiver to measure pseudoranges from no less than four navigation satellites. Interoperability allows receivers to calculate navigation solution in "All-in-view" mode [3], which is making use of all navigation signals from visible satellites.

41.2.1 GDOP Definition

To determine user's coordinates (x, y, z), direction cosine matrix A is used [4]:

$$A = \begin{bmatrix} \cos\alpha_1 & \cos\beta_1 & \cos\gamma_1 & 1 \\ \cos\alpha_2 & \cos\beta_2 & \cos\gamma_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \cos\alpha_N & \cos\beta_N & \cos\gamma_N & 1 \end{bmatrix}. \tag{41.1}$$

where, A is the cosine direction matrix; $\alpha_i, \beta_i, \gamma_i$ are the direction angles between the line of sight from ith satellite and user's location (notice that $\cos^2\alpha_i + \cos^2\beta_i + \cos^2\gamma_i = 1$), N is the visible satellite number. Defined matrix Q as:

$$Q = [A^T \cdot A]^{-1} \tag{41.2}$$

Let σ be the total user equivalent range error (UERE), then we get σ_u, a measure of the overall quality of the least-squares solution, by taking the square root of the sum of the parameter estimate variances:

$$\sigma_u = \sigma\sqrt{Q_{11} + Q_{22} + Q_{33} + Q_{44}}. \tag{41.3}$$

Then we define GDOP as:

$$\text{GDOP} = \sqrt{tr[Q]}, \tag{41.4}$$

where "$tr[]$" is the trace of a matrix, the sum of the elements in the matrix's main diagonal.

41.2.2 GDOP Expectation

41.2.2.1 Calculation

Before study, let's see the following figure first.

Denotes the user's location as the origin point O, built the right hand Cartesian coordination follow East, North and Zenith; denote the satellite location as S. Then the angles α, β, γ are determined as Fig. 41.1. The projection of S in plane xOy is S', the projections of S' in axis x and y are A and B, respectively. Then we obtain:

$$\cos\alpha = \frac{OA}{OS} = \frac{OA}{OS'} \cdot \frac{OS'}{OS} = \cos\angle AOS' \cdot \cos\angle SOS' = \cos\angle AOS' \cdot \sin\gamma,$$
$$\cos\beta = \frac{OB}{OS} = \frac{OB}{OS'} \cdot \frac{OS'}{OS} = \cos\angle BOS' \cdot \cos\angle SOS' = \sin\angle AOS' \cdot \sin\gamma.$$

Denote $\angle AOS'$ as $\angle E$ (the angle between East and OS'); thus, the three variables α, β, γ can be translated into two variables $\angle E$ and γ.

In the following discussion, suppose $\angle E$ is an independent random value distributed uniformly in the interval $[-\pi, \pi]$, and γ_i in the interval $[0, \frac{\pi}{2}]$. Then let's take a look at the matrix multiplication, which is a part of the definition of matrix Q:

$$A^T A = \begin{bmatrix} \cos\alpha_1 & \cos\alpha_2 & \cdots & \cos\alpha_N \\ \cos\beta_1 & \cos\beta_2 & \cdots & \cos\beta_N \\ \cos\gamma_1 & \cos\gamma_1 & \cdots & \cos\gamma_N \\ 1 & 1 & \cdots & 1 \end{bmatrix} \times \begin{bmatrix} \cos\alpha_1 & \cos\beta_1 & \cos\gamma_1 & 1 \\ \cos\alpha_2 & \cos\beta_2 & \cos\gamma_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \cos\alpha_N & \cos\beta_N & \cos\gamma_N & 1 \end{bmatrix}$$
$$= \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \\ d_{31} & d_{32} & d_{33} & d_{34} \\ d_{41} & d_{42} & d_{43} & d_{44} \end{bmatrix} \tag{41.5}$$

where $d_{ij} = d_{ji}$, $(i, j = 1, 2, 3, 4)$, and,

$$d_{11} = \sum_{i=1}^{N} \cos^2\alpha_i = \sum_{i=1}^{N} (\cos^2\angle E_i \cdot \sin^2\gamma_i), \tag{41.6}$$

$$d_{22} = \sum_{i=1}^{N} \cos^2\beta_i = \sum_{i=1}^{N} (\sin^2\angle E_i \cdot \sin^2\gamma_i), \tag{41.7}$$

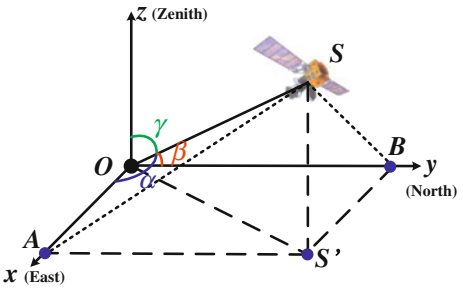


Fig. 41.1 Geometrical description of satellite location

$$d_{33} = \sum_{i=1}^{N} \cos^2 \gamma_i, \tag{41.8}$$

$$d_{44} = N, \tag{41.9}$$

$$d_{12} = d_{21} = \sum_{i=1}^{N} \cos \alpha_i \cos \beta_i = \sum_{i=1}^{N} (\cos \angle E_i \cdot \sin \angle E_i \cdot \sin^2 \gamma_i), \tag{41.10}$$

$$d_{13} = d_{31} = \sum_{i=1}^{N} \cos \alpha_i \cos \gamma_i = \sum_{i=1}^{N} (\cos \angle E_i \cdot \sin \gamma_i \cdot \cos \gamma_i), \tag{41.11}$$

$$d_{14} = d_{41} = \sum_{i=1}^{N} \cos \alpha_i = \sum_{i=1}^{N} (\cos \angle E_i \cdot \sin \gamma_i), \tag{41.12}$$

$$d_{23} = d_{32} = \sum_{i=1}^{N} \cos \beta_i \cos \gamma_i = \sum_{i=1}^{N} (\sin \angle E_i \cdot \sin \gamma_i \cdot \cos \gamma_i), \tag{41.13}$$

$$d_{24} = d_{42} = \sum_{i=1}^{N} \cos \beta_i = \sum_{i=1}^{N} (\sin \angle E_i \cdot \sin \gamma_i), \tag{41.14}$$

$$d_{34} = d_{43} = \sum_{i=1}^{N} \cos \gamma_i. \tag{41.15}$$

Then calculate the expectation of d_{ij}, we obtain:

$$\begin{aligned} E(d_{11}) &= E\left(\sum_{i=1}^{N} \cos \alpha_i^2\right) = E\left(\sum_{i=1}^{N} (\cos^2 \angle E_i \cdot \sin^2 \gamma_i)\right) \\ &= \frac{1}{2\pi} \frac{N}{\frac{\pi}{2}} \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} (\cos^2 \angle E \cdot \sin^2 \gamma) d\gamma d\angle E = \frac{N}{4}, \end{aligned} \tag{41.16}$$

$$\begin{aligned} E(d_{22}) &= E\left(\sum_{i=1}^{N} \cos \beta_i^2\right) = E\left(\sum_{i=1}^{N} (\sin^2 \angle E_i \cdot \sin^2 \gamma_i)\right) \\ &= \frac{1}{2\pi} \frac{N}{\frac{\pi}{2}} \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} (\cos^2 \angle E \cdot \sin^2 \gamma) d\gamma d\angle E = \frac{N}{4}, \end{aligned} \tag{41.17}$$

$$E(d_{33}) = E\left(\sum_{i=1}^{N} \cos \gamma_i^2\right) = \frac{N}{\frac{\pi}{2}} \int_{0}^{\frac{\pi}{2}} \cos^2 \gamma d\gamma = \frac{N}{2}, \tag{41.18}$$

$$E(d_{44}) = N, \tag{41.19}$$

$$\begin{aligned} E(d_{12}) = E(d_{21}) &= E\left(\sum_{i=1}^{N}(\cos\alpha_i\cos\beta_i)\right) \\ &= E\left(\sum_{i=1}^{N}(\cos\angle E_i\cdot\sin\angle E_i\cdot\sin^2\gamma_i)\right), \\ &= \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}}(\cos\angle E\cdot\sin\angle E\cdot\sin^2\gamma)d\gamma d\angle E = 0, \end{aligned} \tag{41.20}$$

$$\begin{aligned} E(d_{13}) = E(d_{31}) &= E\left(\sum_{i=1}^{N}(\cos\alpha_i\cos\gamma_i)\right) \\ &= E\left(\sum_{i=1}^{N}(\cos\angle E_i\cdot\sin\gamma_i\cdot\cos\gamma_i)\right), \\ &= \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}}(\cos\angle E\cdot\sin\gamma\cdot\cos\gamma)d\gamma d\angle E = 0, \end{aligned} \tag{41.21}$$

$$\begin{aligned} E(d_{14}) = E(d_{41}) &= E\left(\sum_{i=1}^{N}(\cos\alpha_i\right) = E\left(\sum_{i=1}^{N}(\cos\angle E_i\cdot\sin\gamma_i)\right) \\ &= \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}}(\cos\angle E\cdot\sin\gamma)d\gamma d\angle E = 0, \end{aligned} \tag{41.22}$$

$$\begin{aligned} E(d_{23}) = E(d_{32}) &= E\left(\sum_{i=1}^{N}(\cos\beta_i\cos\gamma_i)\right) \\ &= E\left(\sum_{i=1}^{N}(\sin\angle E_i\cdot\sin\gamma_i\cdot\cos\gamma_i)\right), \\ &= \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}}(\sin\angle E\cdot\sin\gamma\cdot\cos\gamma)d\gamma d\angle E = 0, \end{aligned} \tag{41.23}$$

$$\begin{aligned} E(d_{24}) = E(d_{42}) &= E\left(\sum_{i=1}^{N}(\cos\beta_i\right) = E\left(\sum_{i=1}^{N}(\sin\angle E_i\cdot\sin\gamma_i)\right) \\ &= \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}}(\sin\angle E\cdot\sin\gamma)d\gamma d\angle E = 0, \end{aligned} \tag{41.24}$$

$$E(d_{34}) = E(d_{43}) = E\left(\sum_{i=1}^{N}(\cos\gamma_i\right) = \frac{N}{\frac{\pi}{2}}\int_{0}^{\frac{\pi}{2}}(\cos\gamma)d\gamma = \frac{2N}{\pi}. \tag{41.25}$$

Thus,

$$
\begin{aligned}
E(Q) &= E\left((A^T A)^{-1}\right) \begin{bmatrix} E(d_{11}) & E(d_{12}) & E(d_{13}) & E(d_{14}) \\ E(d_{21}) & E(d_{22}) & E(d_{23}) & E(d_{24}) \\ E(d_{31}) & E(d_{32}) & E(d_{33}) & E(d_{34}) \\ E(d_{41}) & E(d_{42}) & E(d_{43}) & E(d_{44}) \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \frac{N}{4} & 0 & 0 & 0 \\ 0 & \frac{N}{4} & 0 & 0 \\ 0 & 0 & \frac{N}{2} & \frac{2N}{\pi} \\ 0 & 0 & \frac{2N}{\pi} & N \end{bmatrix}^{-1} = \begin{bmatrix} \frac{N}{4} & 0 & 0 & 0 \\ 0 & \frac{N}{4} & 0 & 0 \\ 0 & 0 & \frac{2\pi^2}{N(\pi^2-8)} & \frac{4\pi}{N(8-\pi^2)} \\ 0 & 0 & \frac{4\pi}{N(8-\pi^2)} & \frac{\pi^2}{N(\pi^2-8)} \end{bmatrix}
\end{aligned} \tag{41.26}
$$

Then, the GDOP expectation should be:

$$
\begin{aligned}
E(\text{GDOP}) &= \sqrt{tr[E(Q)]} = \sqrt{\frac{N}{4} + \frac{N}{4} + \frac{2\pi^2}{N(\pi^2-8)} + \frac{\pi^2}{N(\pi^2-8)}} \\
&= \sqrt{\frac{11 - \pi^2 - 64}{N(\pi^2-8)}} \approx \sqrt{\frac{23.8369}{N}}.
\end{aligned} \tag{41.27}
$$

41.2.2.2 Simulation

To confirm the correctness of formula (41.27), a simulation will be adopted. We will calculate the GDOP value 1,000 times with different satellite number N ($N \in [4, 1000]$). Then statistic the minimal/maximal GDOP value and plot the curves of minimal/maximal GDOP value and formula (41.27). Simulation result shows as Fig. 41.2.

Till now, we obtain the mathematical expectation of GDOP (see expression (41.27)). The variable γ_i was discussed in the interval $\left[0, \frac{\pi}{2}\right]$, which ignores the situation of urban environment where the user is. When the user-satellite link is blocked by buildings in the city, the upper bound of γ_i will no longer be $\frac{\pi}{2}$. The mathematical expectation of GDOP in this case will be discussed in the following section.

41.2.3 GDOP Expectations with Minimal Elevation Changing

41.2.3.1 Calculation

Consider the minimal elevation of user, γ_i should be discussed in the interval $[0, e]$, where e is the complementary angle of minimal elevation of user. Similar with formula (41.16–41.25), we obtain the elements of Q:

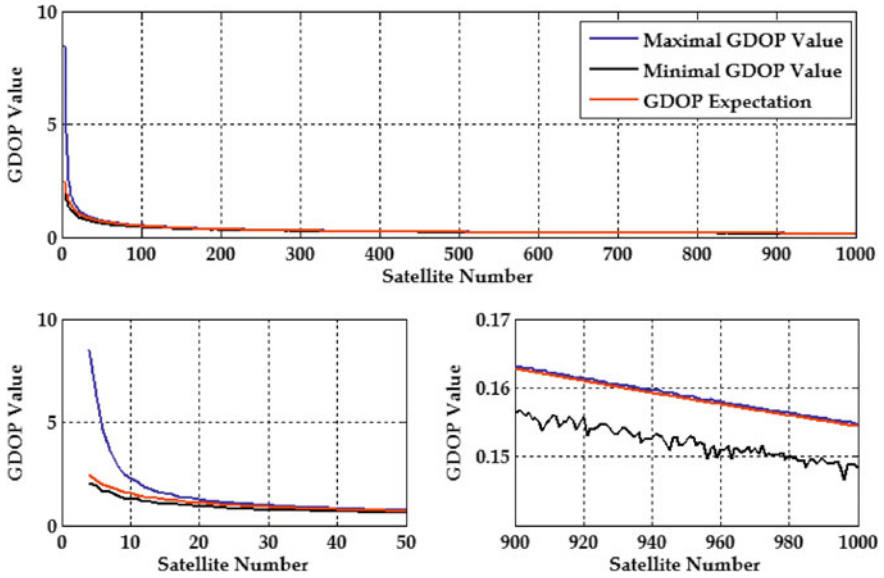


Fig. 41.2 Curves of minimal/maximal GDOP value and formula (41.27)

$$E(d_{11}) = E(d_{22}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\cos^2\angle E\cdot\sin^2\gamma)d\gamma d\angle E = \frac{N(2e-\sin 2e)}{4\pi}, \tag{41.28}$$

$$E(d_{33}) = \frac{N}{\frac{\pi}{2}}\int_{0}^{e}\cos^2\gamma d\gamma = \frac{N(2e+\sin 2e)}{4\pi}, \tag{41.29}$$

$$E(d_{44}) = N, \tag{41.30}$$

$$E(d_{12}) = E(d_{21}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\cos\angle E\cdot\sin\angle E\cdot\sin^2\gamma)d\gamma d\angle E = 0, \tag{41.31}$$

$$E(d_{13}) = E(d_{31}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\cos\angle E\cdot\sin\gamma\cdot\cos\gamma)d\gamma d\angle E = 0, \tag{41.32}$$

$$E(d_{14}) = E(d_{41}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\cos\angle E\cdot\sin\gamma)d\gamma d\angle E = 0, \tag{41.33}$$

$$E(d_{23}) = E(d_{32}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\sin\angle E \cdot \sin\gamma \cdot \cos\gamma)d\gamma d\angle E = 0, \tag{41.34}$$

$$E(d_{24}) = E(d_{42}) = \frac{1}{2\pi}\frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\sin\angle E \cdot \sin\gamma)d\gamma d\angle E = 0, \tag{41.35}$$

$$E(d_{34}) = E(d_{43}) = \frac{N}{\frac{\pi}{2}}\int_{-\pi}^{\pi}\int_{0}^{e}(\cos\gamma)d\gamma = \frac{2N\sin e}{\pi}. \tag{41.36}$$

Thus,

$$\begin{aligned} E(Q) &= E\left((A^TA)^{-1}\right)\begin{bmatrix} E(d_{11}) & E(d_{12}) & E(d_{13}) & E(d_{14}) \\ E(d_{21}) & E(d_{22}) & E(d_{23}) & E(d_{24}) \\ E(d_{31}) & E(d_{32}) & E(d_{33}) & E(d_{34}) \\ E(d_{41}) & E(d_{42}) & E(d_{43}) & E(d_{44}) \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \frac{N(2e-\sin 2e)}{4\pi} & 0 & 0 & 0 \\ 0 & \frac{N(2e-\sin 2e)}{4\pi} & 0 & 0 \\ 0 & 0 & \frac{N(2e+\sin 2e)}{4\pi} & \frac{2N\sin e}{\pi} \\ 0 & 0 & \frac{2N\sin e}{\pi} & N \end{bmatrix} \\ &= \begin{bmatrix} \frac{N(2e-\sin 2e)}{4\pi} & 0 & 0 & 0 \\ 0 & \frac{N(2e-\sin 2e)}{4\pi} & 0 & 0 \\ 0 & 0 & \frac{\pi^2}{N\Lambda} & -\frac{2\pi\sin e}{N\Lambda} \\ 0 & 0 & -\frac{2\pi\sin e}{N\Lambda} & \frac{\pi(2e+\sin 2e)}{2N\Lambda} \end{bmatrix} \end{aligned} \tag{41.37}$$

where,

$$\Lambda = \pi e + \pi\sin e\cos e - 4\sin^2 e \tag{41.38}$$

Then, the GDOP expectation should be:

$$\begin{aligned} E(\mathrm{GDOP}) &= \sqrt{tr[E(Q)]} \\ &= \sqrt{\frac{4\pi}{N(2e-\sin 2e)} + \frac{4\pi}{N(2e-\sin 2e)} + \frac{\pi^2}{N\Lambda} + \frac{\pi(2e+\sin 2e)}{2N\Lambda}} \\ &= \sqrt{\frac{16\pi\Lambda + \pi(2e-\sin 2e)(2\pi+2e+\sin 2e)}{2N\Lambda(2e-\sin 2e)}} \end{aligned} \tag{41.39}$$

It is not clearly enough to see the relationship between GDOP value with satellite number and minimal elevation. So, in the next step, a detailed discussion of the expression (41.39) will be hold.

41.2.3.2 Further Discussion on (41.39)

In this section, the discussion will based on a simulation result, and in the simulation result, we found some interesting phenomena [5–8].

Simulation

To investigate the function character of expression (41.39), the curve will be simulated. Suppose the minimal elevation changes in the interval $[0, \frac{\pi}{2}]$ (i.e. $[0^\circ, 60^\circ]$), the visible satellite number we choose are 6, 8, 10, 15, 30 and 50. Simulation results are shown as Fig. 41.3.

It is clearly that the expectation of GDOP value is an increase function of variable e; but not strict increase. Noticing the six red-dashed-boxes appear in Fig. 41.3, the increase trend of each curve will be mild at some points in these boxes. That is to say, there may be stagnation point in (41.39). Figure 41.4 will show this phenomenon clearly.

In Fig. 41.4, the "mild increasing" trend are more obviously. Thus, we will solve the stagnation point of (41.39).

Inflexion of (41.39)

The derivative of $E(\mathrm{GDOP})$ is:

$$\begin{aligned}\frac{dE(GDOP)}{de} = \sqrt{\pi}\big[&-10\left(64+\pi^2\right)+352\pi e-16\left(2+5\pi^2\right)e^2\\&+\left(960-512\pi e+\pi^2\left(3+48e^2\right)\right)\cos 2e\\&+4\left(43\pi+3e-12\pi^2 e+16\pi e^2+16e^3\right)\sin 2e\\&+2\left(-192+5\pi^2+80\pi e+16e^2\right)\cos 4e\\&+8\pi(-16+56\pi e)\sin 4e+4(7\pi-e)\sin 6e\\&+\left(64-3\pi^2\right)\cos 6e\big]\\&\times\Big[\left(2+\pi\sin 2e-8\sin^2 e\right)^2\\&\cdot 2N(2e-\sin 2e)^2\cdot E(GDOP)\Big]^{-1}\end{aligned} \tag{41.40}$$

Because the denominator of (41.40) is positive, so the sign of (41.40) is dependf on its numerator. The curve in Fig. 41.5 shows that the numerator in $e \in \left[0, \frac{\pi}{2}\right]$ is negative.

Thus, $E(\mathrm{GDOP})$ is always a decreasing function of minimal elevation e. From Fig. 14.4, we also find there may be am inflexion in (41.39). To solve this problem,

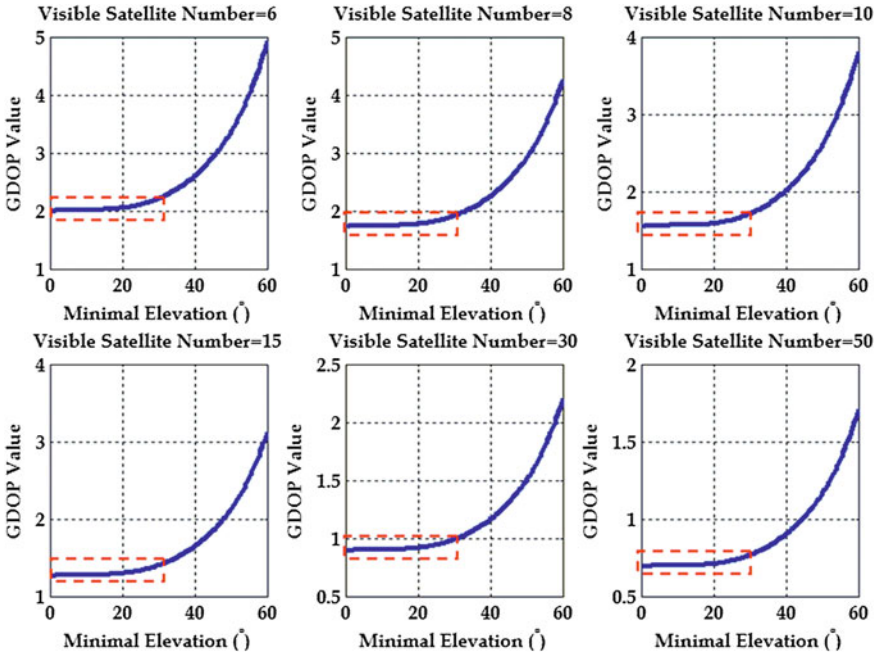


Fig. 41.3 Curves of expectation GDOP value with minimal elevation changing at different visible satellite number

the second derivative of $E(\mathrm{GDOP})$ is required. Because the expression of the second derivative of $E(\mathrm{GDOP})$ is very complex, so we just give the flowing expression:

$$\frac{d^2E(\mathrm{GDOP})}{de^2}\Big|_{e\approx 80.5446^\circ} = 0. \tag{41.41}$$

That is to say, $E(\mathrm{GDOP})$ meets the inflexion point at $e \approx 80.5446^\circ$ (minimal elevation $\approx 9.45544^\circ$). Also the monotonicity of a function does not change at the inflexion point, but the connotation of monotonicity will be change. Thus, we have the conclusion that the GDOP value "gets worse" slowly when minimal elevation less then 9.45544°; but "rapidly gets worse" if the minimal elevation exceed 9.45544°.

41.3 GDOP Expectation in GNSS Constellation

In Sect. 41.2.2, the visible satellite number is hypothesized as a given number N; and we may pay more interests on the GDOP expectation with variable N_T, the total satellite number in a given GNSS constellation.

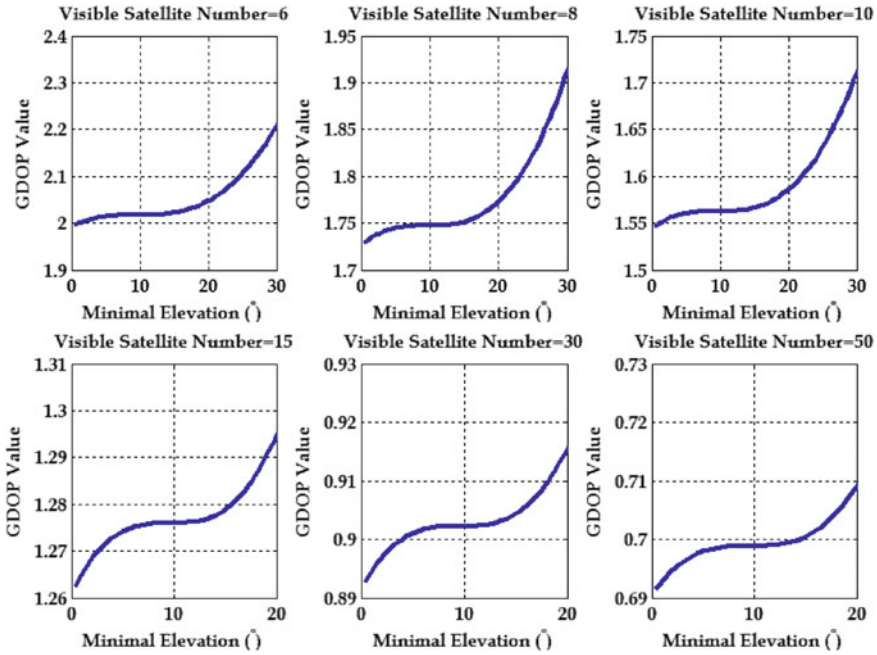


Fig. 41.4 Amplified curves of expectation GDOP value with minimal elevation changing at different visible satellite number

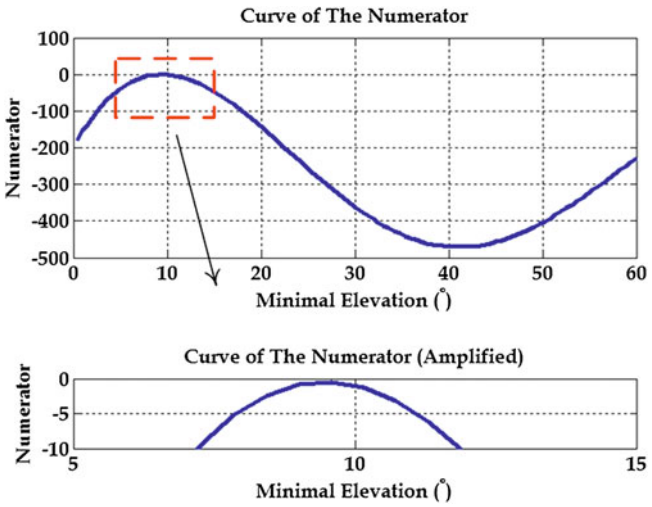


Fig. 41.5 Curve of the numerator in (41.40)

Also in Sect. 41.2.3, the visible satellite number is hypothesized a given number N with the minimal elevation changing. In fact, the visible satellite number gets

smaller as the minimal elevation increasing in a given GNSS constellation. So, there is relationship between visible satellite number N and both the minimal elevation and the total satellite number N_T.

In this section, we will calculate the GDOP expectation in a given GNSS constellation that considering the above discussions.

41.3.1 GDOP Expectation in GNSS Constellation without Minimal Elevation

In this subsection, we will transform the variable N into N_T in expression (41.27) (see Fig. 41.6).

It is clear that the ration from visible satellite number N to total satellite number N_T equals to the ratio from the area of the blue spherical crown to the area of satellite orbit spherical surface, i.e.:

$$\begin{aligned} \frac{N}{N_T} &= \frac{2\pi(R+H)H}{4\pi(R+H)^2} = \frac{H}{2(R+H)} \\ \Rightarrow N &= \frac{N_T H}{2(R+H)} \end{aligned} \tag{41.42}$$

Substitute (41.42) into (41.27), we obtain:

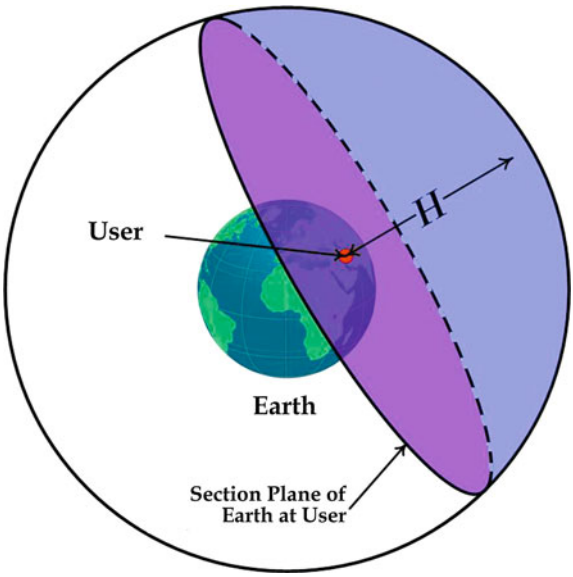


Fig. 41.6 User visible ceiling area and satellite orbit spherical surface

$$E(\mathrm{GDOP}) = \sqrt{\frac{11\pi^2 - 64}{\frac{N_T H}{2(R+H)}(\pi^2 - 8)}} \approx \sqrt{\frac{47.6738(R+H)}{N_T H}}. \tag{41.43}$$

where, H is the satellite orbit height, R is the average earth radius.

41.3.2 GDOP Expectation in GNSS Constellation with Minimal Elevation

Similar as Sect. 41.3.1, we will discuss the GDOP expectation in GNSS constellation with minimal elevation in a sketch map, shown as Fig. 41.7.

It is clear that the ration from visible satellite number N to total satellite number N_T equals to the ratio from the area of the blue spherical crown to the area of satellite orbit spherical surface. The key to calculate this area ratio is to calculate the height of the blue spherical crown (length of BD in Fig. 41.8).

Suppose point P is the user location, there is Cosine Theorem at $\angle APO$ in ΔAPO:

$$\cos\angle APO = \frac{AP^2 + OP^2 - AO^2}{2AP \cdot PO} = -\cos e. \tag{41.44}$$

In $Rt\Delta ABF$, we have:

$$\cos e \frac{BP}{AP}. \tag{41.45}$$

Notice that:

$$OP = R, \quad OA = R + H, \quad DB = H - BP. \tag{41.46}$$

Solve Eqs. (41.3–41.46), we obtain:

$$DB = H - \cos e(\sqrt{H(H+2R) + R^2\cos^2 e} - R\cos e). \tag{41.47}$$

Thus:

$$\begin{aligned} \frac{N}{N_T} &= \frac{2\pi(R+H)DB}{4\pi(R+H)^3} \\ &= \frac{H - \cos e(\sqrt{H(H+2R) + R^2\cos^2 e - R\cos e})}{2(R+H)} \\ \Rightarrow N &= \frac{N_T\left[H - \cos e(\sqrt{H(H+2R) + R^2\cos^2 e} - R\cos e)\right]}{2(R+H)} \end{aligned} \tag{41.48}$$

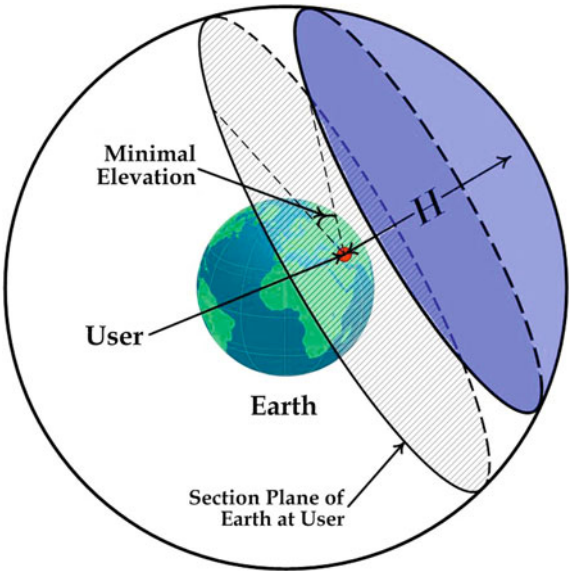


Fig. 41.7 User visible ceiling area and satellite orbit spherical surface considering minimal elevation

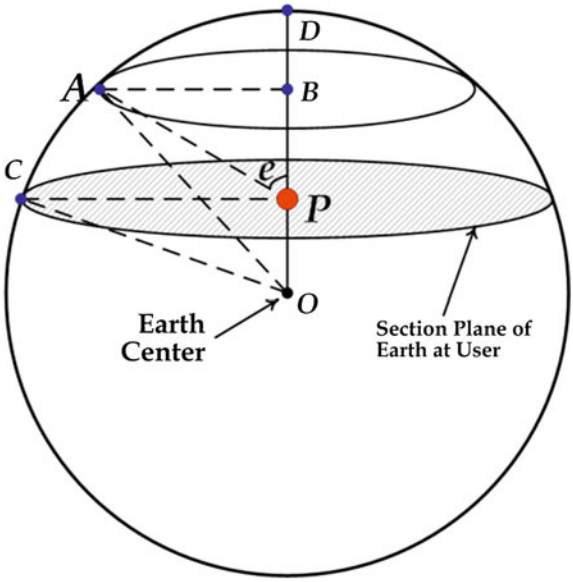


Fig. 41.8 Geometrical sketch map of Fig. 41.7

Submit (41.48) into (41.39), the GDOP expectation in GNSS constellation with minimal elevation is acquired.

41.4 GDOP Expectation in Multi-GNSS Constellation

As GNSS interoperability being a hot topic of GNSS, research on the GDOP in multi-GNSS constellation is also required. In Multi-GNSS constellation,

41.4.1 GDOP Expectation in Multi-GNSS Constellation Without Minimal Elevation

In multi-GNSS constellation, the satellite orbits of different GNSS are different. Suppose the orbit of these satellites are in the interval [*h*,*H*], so the ration from visible satellite number N to total satellite number N_T will be the ration from the volume of blue hollow spherical segment to the volume of hollow sphere (see Fig. 41.9).

Thus, we obtain:

$$\begin{aligned} \frac{N}{N_T} &= \frac{\pi[H^2(R+H-\frac{H}{3}] - h^2(R+h-\frac{h}{3})]}{\frac{4}{3}\pi[R+H)^3 - (R+h^3)]} \\ \Rightarrow N &= 3N_T[H^2(R+\frac{2H}{3}] - h^2(R+\frac{2h}{3})]4[R+H)^3 - (R+h^3)] \end{aligned} \tag{41.49}$$

Submit (41.49) into (41.27), we will obtain the GDOP expectation in multi-GNSS constellation.

41.4.2 GDOP Expectation in Multi-GNSS Constellation with Minimal Elevation

While considering the minimal elevation, the visible satellites are distributing in the blue polyhedron shown as Fig. 41.10.

The volume of the blue polyhedron in the above figure is the key to transform N into N_T. To clarify the construction of the blue polyhedron, the longitudinal section at user's location is employed (Fig. 41.11a); and a coordinate system is built (Fig. 41.11b). Thus, the blue polyhedron is the difference between the geometric solid formed by the red polygon (Fig. 41.11b) rotating around the X axis for a week and the geometric solid formed by the blue polygon (Fig. 41.11b) rotating around the X axis for a week.

Passing over the complex calculation process, then V_0, the volume of the blue polyhedron, is:

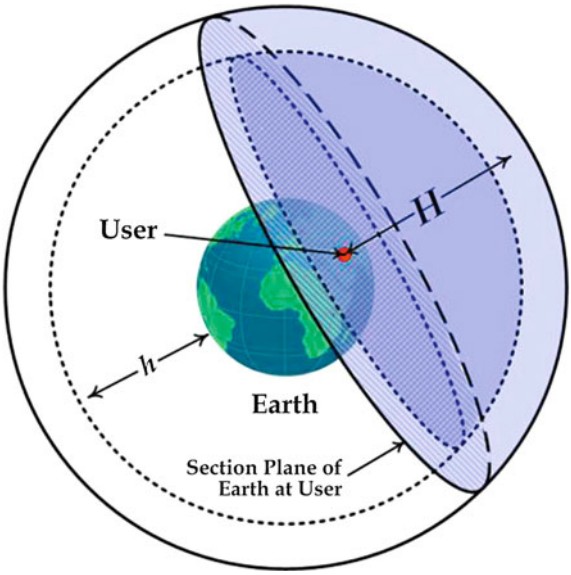


Fig. 41.9 User visible ceiling area and satellite orbit spherical volume

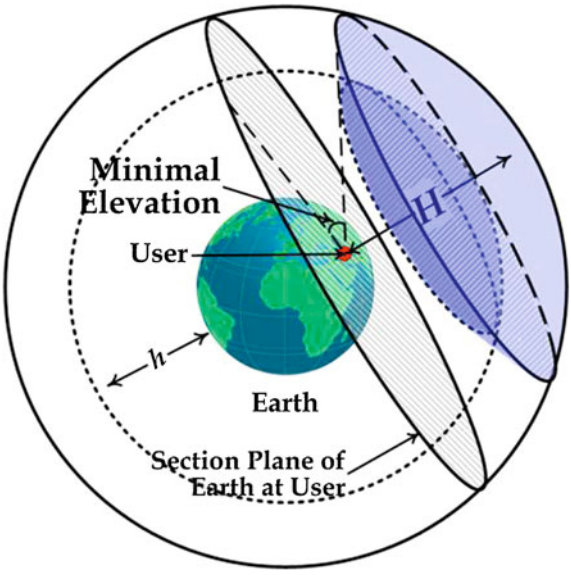


Fig. 41.10 Visible satellite distribution

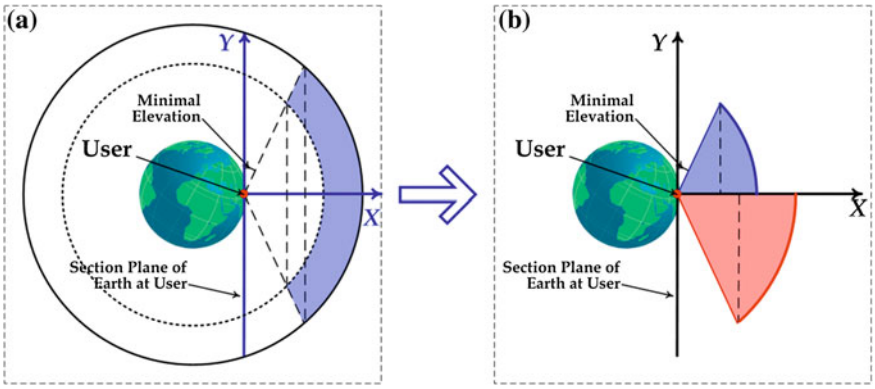


Fig. 41.11 Longitudinal section at user's location (a) and the coordinate system construction (b)

$$\begin{aligned} V_0 = & \frac{1}{3}\pi\cos e\Big\{6(H+R)(\sqrt{H(H+2R)+R^2\cos^2 e}-R\cos e) \\ & +6(h+R)(\sqrt{h(h+2R)+R^2\cos^2 e}-R\cos e) \\ & +\sin e\cos e\Big[(\sqrt{H(H+2R)+R^2\cos^2 e}-R\cos e)^3 \\ & +\sqrt{h(h+2R)+R^2\cos^2 e}-R\cos e)^3 \end{aligned} \tag{41.50}$$

Then we have:

$$\begin{aligned} \frac{N}{N_T} &= \frac{V_0}{\frac{4}{3}\pi[(R+H)^3-(R+h)^3]} \\ \Rightarrow N &= \frac{3N_T V_0}{4\pi[(R+H)^3-(R+h)^3]} \end{aligned} \tag{41.51}$$

Submit (41.51) into (41.39), the GDOP expectation in multi-GNSS constellation with minimal elevation is obtained.

41.5 Application of GDOP Expectation

In this section, we will discuss utilize GDOP expectation in GNSS interoperability and in GNSS constellation evaluation [9, 10].

Table 41.1 Minimal visible satellite number satisfying "GDOP < 1"

Elevation	Minimal visible satellite number for GDOP < 1	
	Calculation result	Round-up to the closest integer
0°	23.8369	24
5°	24.3562	25
10°	24.4213	25
15°	24.5347	25
20°	25.1428	26
30°	29.2590	30
45°	50.8546	51
60°	145.143	146

Table 41.2 Example of using CSI to evaluate constellation stability

Location	CSI of Different Constellations		
	GPS	Beidou	GPS + Beidou
Lintong	0.02816	0.02666	0.00344
Beijing	0.05480	0.01179	0.00340
Sanya	0.02063	0.01151	0.00189
Tokyo	0.04801	0.03046	0.00821
Singapore	0.01410	0.01578	0.00131
New York	0.02135	0.03538	0.00418
Los Angeles	0.06526	0.02365	0.00805
Hawaii	0.03464	0.04913	0.00376

41.5.1 GDOP Expectation in GNSS Interoperability

As we know GNSS interoperability will provides more visible satellite numbers for users and make a better GDOP value. GDOP can be seen as the error "magnification factor", when GDOP value less than 1 it will eliminate the positioning error. By solving the following equation:

$$\frac{16\pi\Lambda + \pi(2e - \sin 2e)(2\pi + 2e + \sin 2e)}{2N\Lambda(2e - \sin 2e)} = 1. \tag{41.52}$$

We obtain:

$$N_0 = \frac{16\pi\Lambda + \pi(2e - \sin 2e)(2\pi + 2e + \sin 2e)}{2\Lambda(2e - \sin 2e)}. \tag{41.53}$$

Then we have Table 41.1.

41.5.2 GDOP Expectation in Constellation Evaluation

In a given constellation, the visible satellite number changes from time to time, the differences between the actual GDOP and the GDOP expectation reflects the stability of the constellation.

The visible satellite number of a given constellation in a period can be determined by simulation process or actual measurement. The square of the difference between the actual GDOP and GDOP expectation with the same visible satellite number can be employed to evaluate constellation performance. Constellation stability index (CSI) can be obtained via Eq. (41.54):

$$\mathrm{CSI} = \mathrm{mean}(\mathrm{GDOP}_{\mathrm{actual}} - E(\mathrm{GDOP}))^2. \tag{41.54}$$

For just give an example, we do not consider the minimal elevation [use Eq. (41.27)]. The constellations we consider are the present GPS constellation, future Beidou constellation and interoperability constellation of GPS and Beidou. Suppose the use is located at Lintong, Beijing, Sanya, Tokyo, Singapore, New York, Los Angeles and Hawaii. The simulation results are shown in Table 41.2.

From Table 41.2, we obtain that the CSI of Beidou is better than the CSI of GPS in Asia area; while the result are opposite in the North America; while the CSI of GPS + Beidou interoperability constellation is the best.

41.6 Conclusion

In this paper, we calculated the formula of mathematical GDOP expectation. The matrix elements for GDOP calculation are simplified by geometrical analyses. After the complexity calculation, we discussed the GDOP expectation in GNSS constellation and Multi-GNSS constellation. At last, the applications of GDOP expectation in interoperability and constellation stability evaluation are introduced; examples are given in each application area.

References

1. Providers' Forum Principles of Compatibility and Interoperability and their further definition (2012) International Committee on GNSS, Providers' Forum. http://www.oosa.unvienna.org. Accessed 11 Dec 2012
2. Report of the First Meeting of the International Committee on Global Navigation Satellite Systems (ICG) (2007) Providers forum. In: The first meeting of ICG, Bangalore, 4 Sept 2007
3. Stupak G (2008) The Russian federation view on GNSS compatibility and interoperability. In the third meeting of the international committee on global navigation satellite systems (ICG), Pasadena, USA, 8–12 Dec
4. Langley RB (1999) Dilution of precision. GPS World 10:52–59

5. Han T, Lu X et al (2011) Differential equation dynamical system based assessment model in GNSS interoperability. Sci China: Phys Mech Astron 54(6):996–1003
6. Lu J, Yang Q (2010) Study on GNSS interoperability. Sci China: Phys Mech Astron 40(5):1–8
7. Li J, Li Z, Hao J et al (2009) A preliminary study on compatibility and interoperability of GNSS. J Geometrics Sci Technol 26(3):177–180
8. Han T, Lu X, Rao Y et al. (2011) Saturated GDOP value analysis in satellite navigation constellation. In: CSNC-2011, Shanghai
9. Lu X, Lu J, Bai Y et al. (2010) Interoperability feasibility analysis between beidou and GPS. In: Global navigation satellite systems: report of a joint workshop of the National Academy of Engineering and the Chinese Academy of Engineering, pp 75–82
10. Wang H, Zhang X, Zhang Y (2008) Geometric dilution of precision for GPS single-point positioning based on four satellites. J Syst Eng Electron 19(5):1058–1063

Chapter 42
Real-Time Coseismic Velocity and Displacements Retrieving and De-Noising Process by High-Rate GNSS

Rui Tu, Rongjiang Wang, Yong Zhang, Maorong Ge and Qin Zhang

Abstract High-rate GNSS is commonly used in seismology, as the impacts of orbit error, clock error, atmosphere error and multipath effect, the coseismic velocity and displacement which are estimated by the broadcast ephemeris usually exists drifts, and serious distort by the pollution of high-rate noise. This paper proposes a new method with the low pass filter and linear correction are used to de-trending low frequency drifts and de-noising high frequency noise by S transform, then the coseismic velocity and displacement are real-time recovered. The validations show that, the recovered velocity is very consistent with the waveforms recorded by strong-motion, the precision of coseismic displacement is better than 2 and 5 cm in horizontal and vertical respectively, the inversion results which gained by real time recovered displacements is closed to the post processed results.

Keywords Real-time · Coseismic velocity · Coseismic displacement · De-noising · High-rate GNSS

42.1 Introduction

Using high-rate Global Navigation Satellite System (GNSS) technology to research seismology has become one of the most popular topics in GNSS application area [1–8]. Currently, the network solution and Precise Point Positioning (PPP) are the mainly two modes for high-rate GNSS data processing, both of them

R. Tu (✉) · R. Wang · Y. Zhang · M. Ge
The Germany Research Center for Geosciences, Telegrafenberg 14473, Potsdam, Germany
e-mail: turui-2004@126.com

Q. Zhang
Chang An University, Yanta road No.126, Xi'An 710054, China

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_42,

need high precision products such as orbit, clock and earth rotation parameters [9]. The release of the high precision products has a long time's delay, meanwhile the real-time products with low precision and bad reliability, it is limited for real-time earthquake monitoring and early warning. What's more, while for network solution, one or more static stations are needed as reference stations, it is difficult to find the reference stations nearby the epicenter as seismic waves sweep hundreds or 1,000 km away. If the reference station is selected far away, the network's precision of dynamic positioning is worse because of the longer baseline, and the network solution can only get relative displacement, it is difficult to describe the activity of earthquake. While for PPP, it is flexible and has no influence with the baseline length, but it with the shortcoming of long convergence, low precision and poor reliability because many errors can not accuracy processed in the Un-difference process model [10]. To solve these problems, Colosimo et al. [11] proposed a new idea to estimate the velocity and recovery the coseismic displacement by using the broadcast ephemeris and a single receiver in real-time. But it causes a new problem that the coseismic velocity and displacement is drifting and noising as the orbit error, clock error, atmosphere error, multipath effect and high frequency noises.

This paper propose a new method for real time retrieving and de-noising coseismic velocity and displacement. The main idea is that the coseismic velocity is estimated based on broadcast ephemeris and the epoch-difference observation model in real-time, then the true coseismic velocity and displacement are retrieved by de-trending and de-noising process. The validation of the real seismic data processing show that the recovered velocity is very consistent with the waveforms recorded by strong-motion, the precision of coseismic displacement is better than 2 and 5 cm in horizontal and vertical direction respectively, the inversion results which gained by real time recovered displacement is consistent to the post processed results.

42.2 The Principle of Velocity Estimation by GNSS

For the carrier phase observations, the velocity estimation model by the epoch-difference can be described as follows [11]:

$$\lambda\Delta\Phi_r^s = (e_r^s \cdot \Delta\xi_r + c\Delta\delta t_r) + \left(\left[\Delta\rho_r^s\right]_{OR} - c\Delta\delta t^s\right) + (\Delta E_r^s + \Delta m_r^s + \Delta\varepsilon_r^s) \quad (42.1)$$

where, λ is the carrier phase wavelength, Δ is epoch-difference, subscript s, r refers to satellite and receiver respectively, Φ is the phase observation, e is the unit vector from satellite to the receiver, ξ is the station position, δt_r is receiver clock error, $\left[\rho_r^s\right]_{OR}$ is the broadcast ephemeris orbit error, δt^s is the satellite clock error, E is the errors which contain troposphere, ionosphere, relativity and phase wind-up, m is the multipath effect, ε is the observation noise.

The stochastic model is as follows:

$$w([\lambda\Delta\Phi_r^s]) = \cos^2(Z) \tag{42.2}$$

Z is the satellite elevation. Based on the mathematical model, we can do least squares solution while more than four satellites are tracked for each continuous epoch. The unknown parameters are velocity in three dimensional and epoch-difference of the receiver clock.

42.3 Coseismic Velocity and Displacements Retrieving and De-Noising Process

In the velocity estimation model, as the orbit error, clock error, atmosphere error and multipath effect, the estimated velocity has a drift, it will be further amplified while integrate to displacement. Meanwhile, the real seismic wave signal is greatly influenced by the noise of the high frequency GNSS. In order to retrieve coseismic displacement and waveform accuracy, it must be do displacement retrieving and signal de-noising process. The method of recovery the coseismic displacement and de-noising high frequency GNSS signal noise is performed in seven steps. Firstly, three successive time windows can be chosen to define the pre-seismic, co-seismic and post-seismic periods, respectively. Secondly, a low pass filter is used to de-trending the drift and get the low frequency information. Thirdly, the low frequency signal is extracted by linear fitting the low frequency information. Fourth, an orthogonal polynomial is used to de-trending the raw velocity and extract the high frequency information. Fifth, the S transform is used to de-noising the high frequency information and get the high frequency signal. Sixth, get the real coseismic velocity signal by the combination of the low frequency signal and high frequency signal. Seventh, integrate the velocity to coseismic displacement.

42.3.1 Determine the Window of Sensitive Signal

In the time series of coseismic velocity, we define an average energy function, like Eq. (42.3).

$$W_n = \frac{1}{N}\left(\sum_{t=n+1}^{n+N} v(t)\cdot\Delta t\right), W_0 \tag{42.3}$$

where, W_n is the average energy at epoch n, W_0 is the rest energy, it reacts the noise level of velocity, N is epoch numbers, $v(t)$ is velocity time series, Δt is sample interval. The signal begin time t_1 is decide while the average energy W_n is more than five times of the rest energy W_0, the end time t_2 is decide while average energy

W_n less than 5 times of the rest energy W_0 and time after the begin time, it shows in Fig. 42.1, then define the pre-seismic, co-seismic and post-seismic periods.

42.3.2 De-Trending the Drift by Low Pass Filter

For the coseismic velocity times series, the low pass filter is used to de-trending the low frequency drift and retrieve the real low frequency information [12]. The time series of GNSS raw velocity is as following.

$$v(t) = (V_1, V_2, V_3 \ldots V_N)^T \tag{42.4}$$

where, N is epoch numbers, v is the raw velocity, it contains the low frequency drift and low frequency information, like Eq. (42.5), $v_{stat}(t)$ is the low frequency information, it contains the low frequency signal and linear shift.

$$v(t) = v_{stat}(t) + v_{trend}(t) \tag{42.5}$$

The drift $v_{trend}(t)$ can be expressed with a linear observation model, like Eq. (42.6).

$$v_{trend}(t) = I\theta + \upsilon \tag{42.6}$$

where, $I \in R^{N\times N}$ is the observation matrix, $\theta \in R^N$ are the regression parameters, υ is the observation error. While using the regularized least squares solution, the results of the low pass filter as follows.

$$\hat{v}_{stat}(t) = v(t) - I\hat{\theta}_\lambda = (I - (I + \lambda^2 D_2^T D_2)^{-1})v(t) \tag{42.7}$$

where, λ is the regularization parameter and $D_2 \in R^{(N-2)\times(N)}$ indicates the discrete approximation of two derivative operator, like Eq. (42.8).

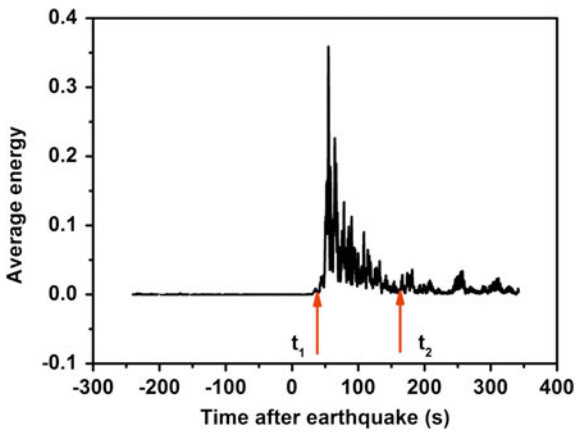


Fig. 42.1 Determine the window of sensitive signal

$$D_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix} \tag{42.8}$$

Figure 42.2 is the comparison of low pass filter velocity and raw velocity.

42.3.3 Extract Low Frequency Signal by Linear Fitting

For the velocity which processed by low pass filter, make a linear fitting with weight of Eq. (42.9), the linear drift correction coefficient can be solved [13].

$$\hat{v}_{stat}(t) = at + b \qquad \begin{cases} p = 0(t_1 < t < t_2) \\ p = 1(t \le t_1, t \ge t_2) \end{cases} \tag{42.9}$$

The low frequency signal which retrieved from de-trending the linear drift can be written like Eq. (42.10), Fig. 42.3 shows the low frequency signal.

$$v_{low}^{offset}(t) = \hat{v}_{stat}(t) - (\hat{a}t + \hat{b}) \tag{42.10}$$

42.3.4 De-Trending by Orthogonal Polynomial

For the raw velocity time series, make a orthogonal polynomial correction with Eqs. (42.11–42.14), the residuals contain high frequency noise and high frequency signal, like Eq. (42.15).

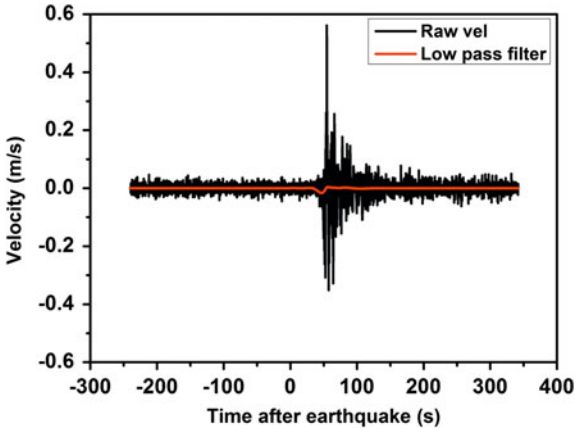


Fig. 42.2 The comparison of the low pass filter velocity, the *black line* is the raw velocity and the *red line* is the low pass filter velocity

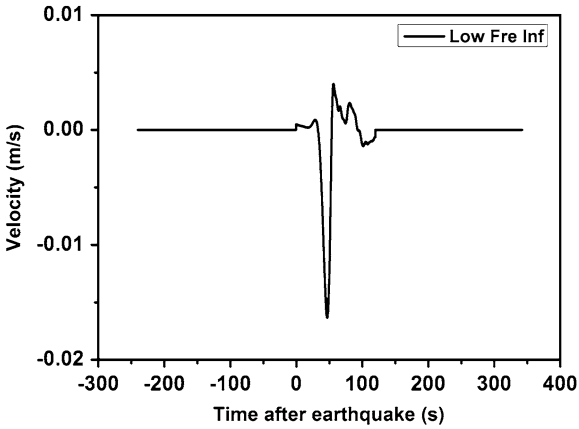


Fig. 42.3 The graph of the low frequency signal

$$v_{fit}(t) = \sum_{i=0}^{N} b_i p_i(t) \tag{42.11}$$

$$\int_0^T p_i(t)\, p_j(t)\, w(t) dt = 0, \quad i \neq j \tag{42.12}$$

$$\int_0^T |p_i(t)|^2 w(t) dt = 1, \quad i = 0, 1, \ldots n \tag{42.13}$$

$$b_i = \int_0^T v(t) p_i(t)\, w(t) dt \tag{42.14}$$

$$v_{high}(t) = v(t) - v_{fit}(t) \tag{42.15}$$

where, $p_i(t)$ is orthogonal function, b_i is orthogonal factor, $w(t)$ is weight function, we treat with equivalent weight is this study. $v_{fit}(t)$ is the polynomial correction, $v_{high}(t)$ is the high frequency information which contains high frequency signal and noise. In Fig. 42.4, it shows the orthogonal polynomial correction curve.

42.3.5 De-Noising by S Transform

For the high frequency part which retrieved in the fourth step, do de-nosing process with S transform [14], Eq. (42.16) shows the S transform.

$$S(\tau, t) = \int_{-\infty}^{\infty} v_{high}(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi f t} dt \tag{42.16}$$

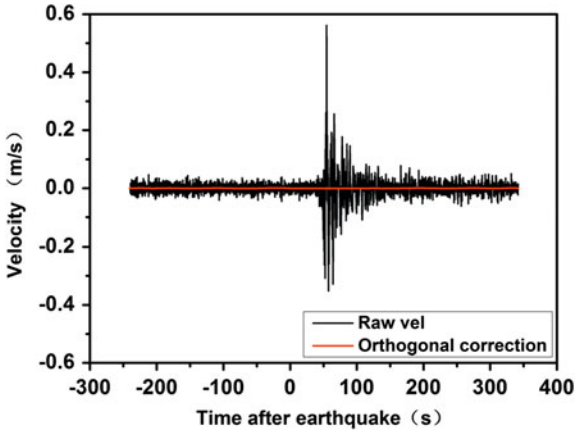


Fig. 42.4 The correction chart of the orthogonal polynomial, the *black line* is the raw velocity and the *red line* is the orthogonal correction velocity

The de-noising processed according to the following threshold criterion in the frequency area.

$$c(x) = \begin{cases} x - \mathrm{sgn}(x)(1-\alpha)\lambda & |x| \geq \lambda \\ 0 & |x| \leq \gamma \\ \alpha\lambda\left(\frac{|x|-\gamma}{\lambda-\gamma}\right)^2\left((\alpha-3)\left(\frac{|x|-\gamma}{\lambda-\gamma}\right)+4-\alpha\right) & others \end{cases} \tag{42.17}$$

$$\lambda = \delta\sqrt{2\log N} \tag{42.18}$$

where, x is either the real or imaginary part of $S(\tau, t)$, the parameters γ, λ, α satisfy these conditions, $0 < \gamma < \lambda, 0 \leq \alpha \leq 1$, in this study, $\gamma = 0.5\lambda$, $\alpha = 0.7$. δ^2 is the frequency-dependent noise variance and N is the length of signal. After de-noised the real part and imaginary part with the threshold, do the S inverse transform like Eq. (42.19) to get the high frequency information. Figure 42.5 shows the comparison of the high frequency information by de-noising.

$$v_{high}^{denoise}(t) = \sqrt{2\pi}\int_{-\infty}^{\infty}\frac{S(t,f)}{|f|}e^{i2\pi ft}df \tag{42.19}$$

42.3.6 Combine of Coseismic Velocity

The real coseismic velocity is the combination of low frequency de-trending information and high frequency de-noising information, like Eq. (42.20) and Fig. 42.6 shows.

$$v_{com}(t) = v_{low}^{deoffset}(t) + v_{high}^{denoise}(t) \tag{42.20}$$

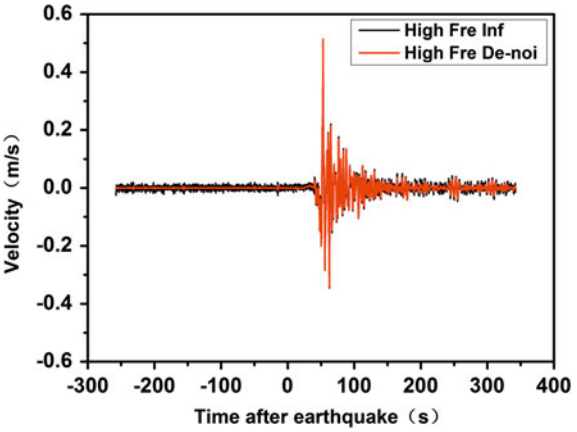


Fig. 42.5 The comparison of high frequency information by de-noising, the *black line* is the high frequency part, it contains the information and signal, the *red line* is the high frequency de-noising information

42.3.7 Retrieving of Coseismic Displacement

The synthesis velocity integral to coseismic displacement with Eq. (42.21). Figure 42.7 shows the low frequency and high frequency displacement informations.

$$s_{com}(t) = \int v_{com}(t)dt \tag{42.21}$$

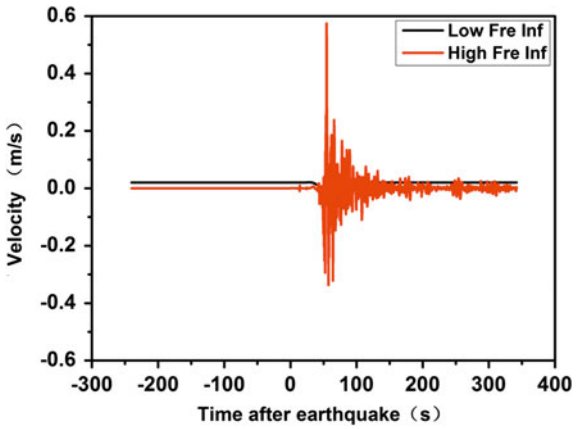


Fig. 42.6 The synthetic map of coseismic velocity, the *black line* is the low frequency velocity information, the *red color* is the high frequency velocity information

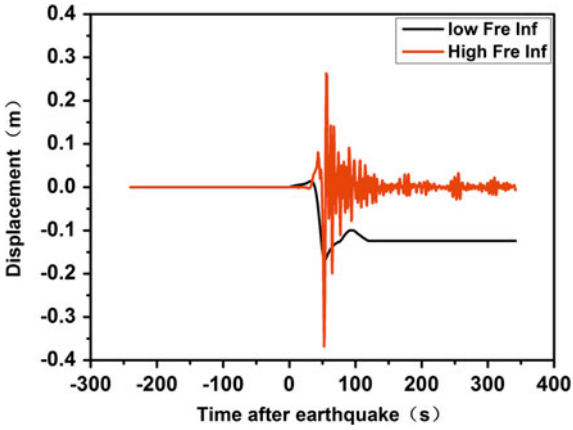


Fig. 42.7 The recovery diagram of coseismic displacement, the *black line* is the low frequency displacement information, the *red color* is the high frequency displacement information

42.4 Validation

42.4.1 Data

In order to validate the effectiveness of the method proposed in this study, the Baja earthquake which occurred in California of Mexico is selected as an example (Epicenter: 2.259°N, 115.287°W, Time: 4 April, 2010, UTC 22:40:42, Magnitude: Mw7.2, Depth: 10 km). 34 continuous GNSS stations nearby the epicenter are collected for real-time velocity estimation and displacement recovery, the sample interval is 5HZ. Figure 42.8 shows the distribution of GNSS stations and epicenter.

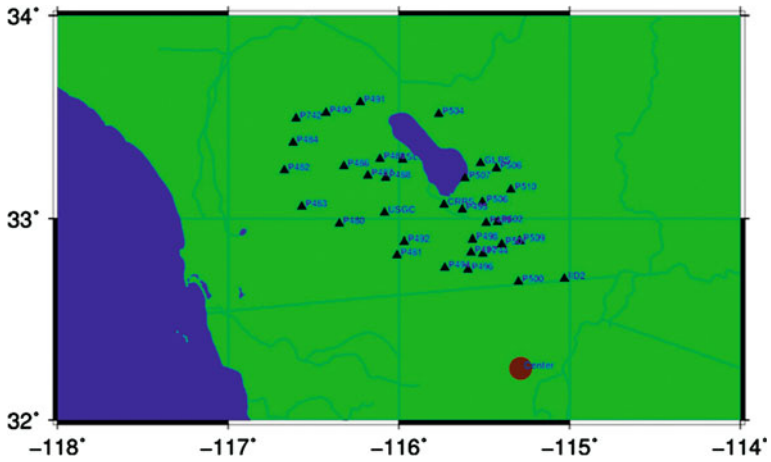


Fig. 42.8 The distribution of epicenter and GNSS stations, the *blue triangle* is the GNSS stations, the big *red circle* is the epicenter

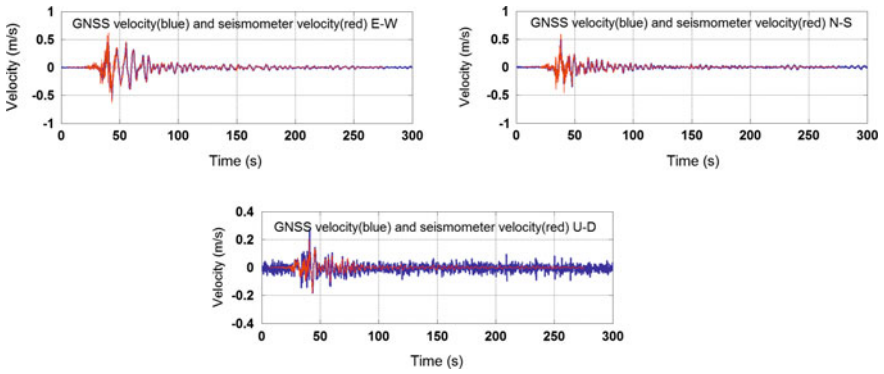


Fig. 42.9 The comparison of velocity between GNSS and strong motion, the *blue line* is the GNSS velocity, the *red line* is the seismometer velocity, from *left* to *right* are the East–West, South-North and *Vertical* direction respectively

42.4.2 The Velocity Compare Between GNSS and Strong Motion

Figure 42.9 show the comparison of velocity which gained from GNSS station P496 and strong motion station 5058, it can be found that the velocity estimated by GNSS is consistent with the waveform recorded by the high sensitivity strong-motion.

42.4.3 De-Trending and De-Noising Analysis of Coseismic Velocity and Displacement

Figure 42.10 is the velocity time series of station P496, it shows that the high frequency GNSS velocity estimation which based on broadcast ephemeris exist obvious noise pollution, the noise level is 0.009, 0.014, 0.049 m/s in ENU direction respectively in this example. The high frequency noise seriously disrupted the authenticity of the seismic signal, especially in weak seismic signal and high noise level condition such as height direction, the seismic information is more difficult to obtain accuracy. Figure 42.11 is the time series process of recovering the permanent displacement. Due to the orbit error, clock error, atmosphere error and multipath effect, the estimated velocity has a drift, it is further amplified while integrate to displacement, as shown in the black line. After de-trending process, the coseismic displacement is recovered, showing in red line. Meanwhile, the real seismic information is retrieved by the high frequency de-noising process, showing in blue line.

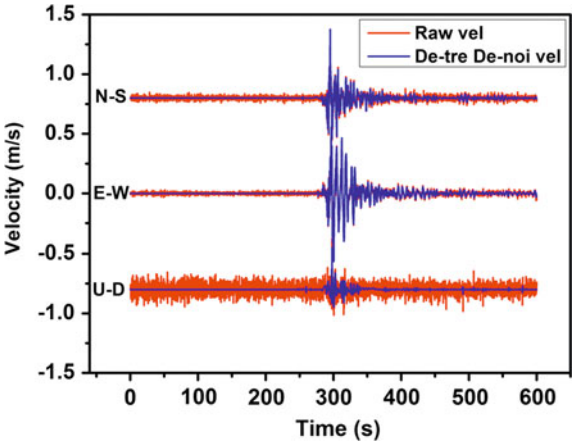


Fig. 42.10 The time series of velocity, the red *line* is the raw velocity, the *blue line* is the velocity after de-trend and de-noise

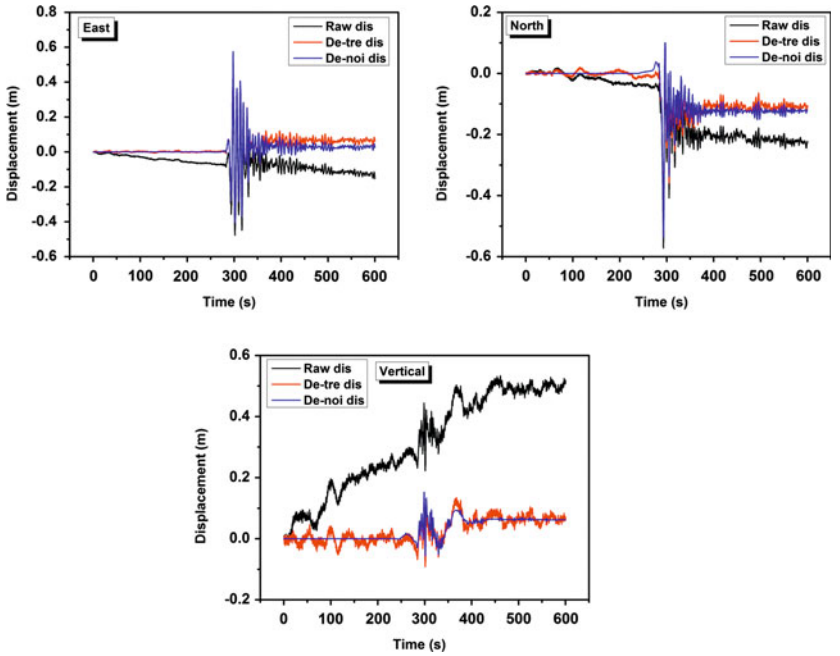


Fig. 42.11 The time series of displacement, the *black line* is the raw displacement, the *red line* is the displacement after de-trend, the *blue line* is the displacement after de-trend and de-noise, from *left* to *right* are the East–West, South-North and *Vertical* direction respectively

42.4.4 The Statistical Analysis and Inversion Interpretation of the Permanent Displacement

In order to analysis the permanent displacement, we using the software Gammit to get the true permanent displacement, where the difference between daily solutions of the day before the earthquake and the day after the earthquake is treat as reference permanent displacement, the precision is better than 5 mm. Figure 42.12 showing the comparison of the coseismic permanent displacement between real time recovered by the new method and the static post network solution. It can be concluded that the coseismic displacement recovered by de-trending and de-noising has a good consistency with the post solution results, the RMS values of residuals are (1.1, 1.4, 2.8) cm in ENU direction respectively.

In order to validate the reliability of the method proposed in this paper further more, we make a comparison between the real-time broadcast ephemeris GNSS process results and final precise ephemeris process results with the finite fault static dislocation inversion. According to the GCMT results by Harvard University, the fault geometric parameters as follows, strike 313°, dip 88°, the rake angle is allowed to vary ±20° around 187°, the fault size is 130 km along the strike and with width 20 km, which are divided into 26*4 sub-faults, each sub-faults size is

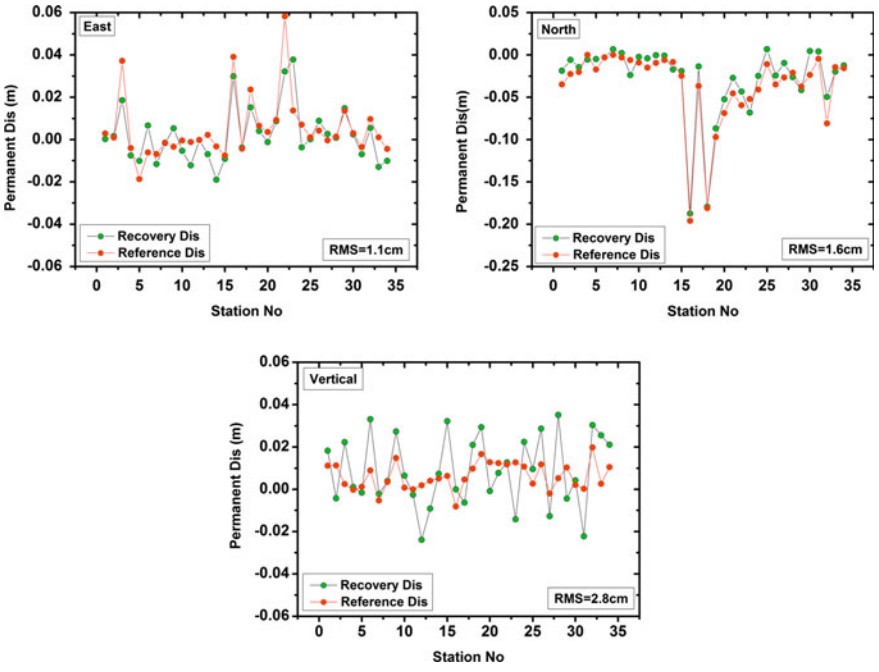


Fig. 42.12 The compare of permanent displacement in East–West, South-North and *Vertical* direction, the *red circle* is the reference displacement, the *green circle* is the recovered displacement

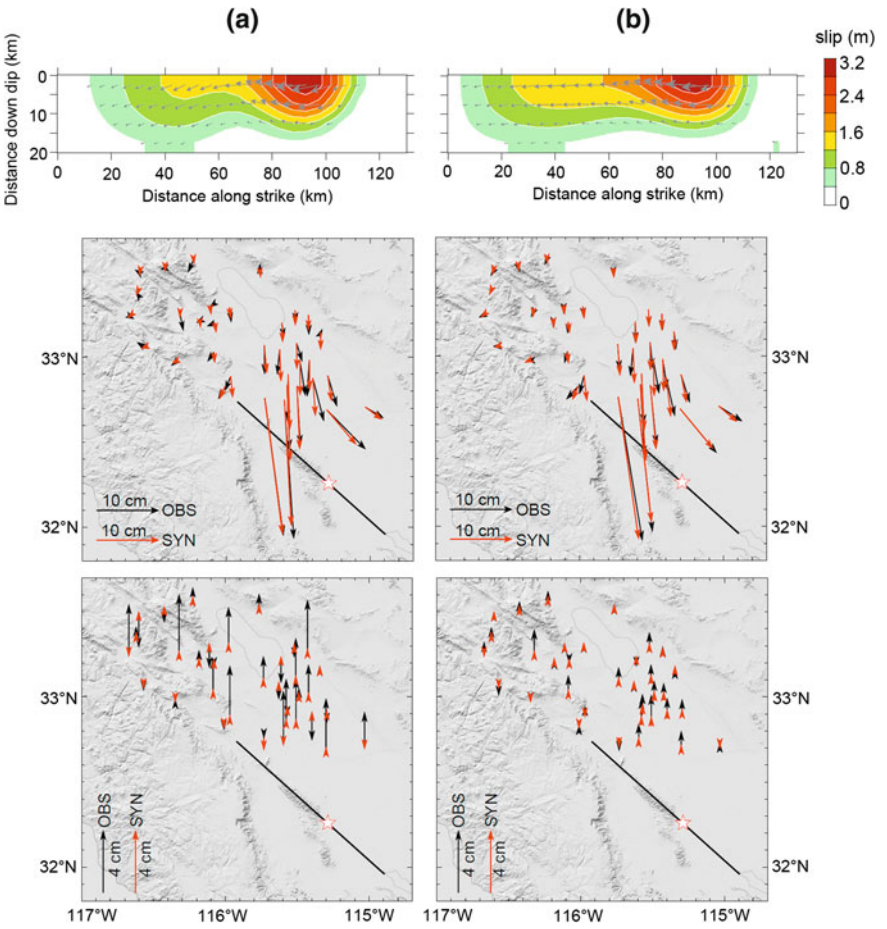


Fig. 42.13 The static sliding momentum distribution inversed by different coseismic displacement. **a** Broadcast ephemeris real time processing results. **b** Final precise ephemeris processing results. From *top* to *bottom* are static sliding momentum distribution, the compare of the observation displacement and synthetic coseismic displacement in *horizontal* direction and *vertical* direction. The *red star* is the epicenter, the *black thick line* is the intersecting line of the earthquake fault and the ground

5*5 km. The inversion programme which developed by Wang et al. [15] is used to analysis and compare the static sliding distribution of the fault.

The moment magnitude inversed by the real-time results and post results are nearly the same, they are Mw7.12 and Mw7.17 respectively. As showing in Fig. 42.13, the sliding distribution of two results with good consistent, the maximum sliding momentum is 3.14 m, it is located along the strike of 80–100 km in the fault plane direction. The fracture area of the fault is mainly concentrated in the near-surface, and the fracture is obviously outcropping. Except the magnitude difference is about 0.05 Mw for the two inversion results, there also some other differences,

such as the extension range of fault movement in the strike direction is smaller by the real-time results. The differences are the reflection of different coseismic displacement, while the precision is very similar with the offset 1–2 cm in the horizontal direction, we can speculated that the main reason is the difference of the vertical direction, as in vertical direction, the errors and noise is more difficult to accuracy process. Even so, the real-time results processed by broadcast ephemeris can still obtain the effective information of source fracture quickly and accurately. It will be used for real-time estimation of earthquake magnitude, rupture scale and scope, and other important source parameters, meanwhile the first-hand material about the source information for the earthquake disaster estimation can also be provided.

The magnitude which inversed by the real-time results has a good consistency with some other researches in the main source information, it is a little bigger than the Allen and Zin [16]'s results, but very close to the GCMT result and Crowell et al. [17]'s results. The fault slip momentum distribution which inversed whether by real-time process results or post process results also very similar to Crowell et al. [17]'s results. Both of them reflect the source fracture characteristics that the fracture occurred in near-surface and the fracture outcropping. This implies that while using the broadcast ephemeris to real-time estimate the velocity and recovery the displacement, it can also analysis and evaluate the main source parameters accuracy, this will greatly improve the response speed after the earthquake compared to the final precise ephemeris processing.

42.5 Conclusion

In the area of GNSS seismology, as the residual errors and high frequency noise, using broadcast ephemeris to real-time estimate velocity and recover displacement exist drift and distortion phenomenon. In this study, we propose a new method for real-time coseismic velocity and displacements retrieving and de-noising process by high-rate GNSS. It using a low pass filter and linear correction to de-trend low-frequency drifts and de-noise high frequency noise by S transform, then the real seismic informations are real-time recovered. The validation results of the Baja earthquake in Mexico show that, the broadcast ephemeris real-time estimated velocity is consistent with the waveforms recorded by the strong-motion. The coseismic displacement after de-trending and de-noising process is anastomosis with the post network solution results, with precision 2 cm and 5 cm in horizontal and vertical direction respectively, the coseismic displacement inversion results is very close between real-time recovered and post processed.

In the future, with the development of high frequency GNSS and construction of multimode satellite navigation system, single frequency receiver and broadcast ephemeris can be used for real time velocity estimating and coseismic displacement recovering, meanwhile with the advantages of high precision, low cost and convenient, it may be widely used in real-time earthquake monitoring and early warning.

Acknowledgments We recognize the fundamental role of the IGS and the California Real time network for delivering high-rate GNSS data, we also thanks doctor Guo Aizhi's guidance communication.

References

1. Blewitt G, Kreemer C, Hammond WHP, Stein S, Okal E (2006) Rapid determination of earthquake magnitude using GPS for tsunami warning systems. Geophys Res Lett 33:L11309. doi:10.1029/2006GL026145
2. Blewitt G, Hannond W, Kreemer C, Plag H, Stein S, Okal E (2009) GPS for real-time earthquake source determination and tsunami warning system. J Geod 83:335–343. doi:10.1007/s00190-008-0262-5
3. Bock Y, Nikolaidis R, Jonge P (2000) Instantaneous geodetic positioning at medium distances with the Global Positioning System. J Geophys Res 105(B12):28, 223–28, 253
4. Genrich J, Bock Y (2006) Instantaneous geodetic positioning with 10–50 Hz GPS measurements: noise characteristics and implications for monitoring networks. J Geophys Res 111:B03403. doi:10.1029/2005JB003617
5. Bock Y, Prawirodirdjo L, TI Melbourne T (2004) Detection of arbitrarily large dynamic ground motions with a dense high rate GPS network. Geophys Res Lett 31:L06604
6. Choi K, Bilich A, Larson K, Axelrad P (2004) Modified sidereal filtering: implications for high rate GPS positioning. Geophys Res Lett 31:L22608
7. Larson K, Miyazaki S (2008) Resolving static offsets from high-rate GPS data: the 2003 Tokachi-oko earthquake. Earth Planets Space 60:801–808
8. Larson K, Bodin P, Gomberg J (2003) Using 1 Hz GPS data to measure deformations caused by the Denali Fault earthquake. Science 300(5624). doi:10.1126/science.1084531
9. Dow JM, Neilan R, Rizos C (2009) The international GNSS service in changing landscape of global navigation satellite systems. J Geod 83:191–198
10. Fang R (2010) High-rate GPS data Non-difference precise processing and its application on seismology. Dissertation, Wuhan University
11. Colosimo G, Crespi M, Mazzoni A (2011) Real-time GPS seismology with a stand-alone receiver: a preliminary feasibility demonstration. J Geophys Res 116:B11302
12. Tarvainen MP, Ranta-aho PO, Karjalainen PA (2002) An advanced detrending method with application to HRV analysis. IEEE Trans Biomed Eng 49(2):172–175
13. Zhu L (2003) Recovering permanent displacements from seismic records of the June 9, 1994 Bolivia deep earthquake. Geophys Res Lett 30:1740
14. Parolai S (2009) De-noising of seismograms using the S-transform. Bull Seism Soc Am 99(1):226–234
15. Wang R, Lorenzo-Martin F, Roth F (2003) Computation of deformation induced by earthquakes in a multi-layered elastic crust-Fortran programs EDGRN/EDCMP. Comp Geosc 29:195–207
16. Allen R, Ziv A (2011) Application of real-time GPS to earthquake early warning. Geophys Res Lett 38(16):L16310. doi:10.1029/2011GL047947
17. Crowell BW, Bock Y, Melgar D (2012) Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment. Geophys Res Lett 39:L09305. doi: 10.1029/2012GL051318

Chapter 43
Analysis of the Global Water Vapor Distribution with COSMIC Radio Occultation Observations

Jia Luo, Zhiping Chen and Lei Yi

Abstract Water vapor is one of the most important greenhouse gases in the Earth's atmosphere. It plays a crucial role in the energy and hydrological cycles of the Earth's climate system. In this study, we compare the wet profiles derived from Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) radio occultation (RO) during 2009 with those from radiosondes. The results show that the specific humidity profiles derived from COSMIC RO observations agree well with those from radiosonde observations. In addition, the global distribution of water vapor derived from COSMIC RO shows that the gradient variation from low latitude to high latitude is distinct except in some special areas, such as the Sahara desert of North Africa, which reflects the impacts of upward and downward branches of the Hardley cell, Ferrel cell and the polar circulation during NH summer and NH winter. The contrast between dry and moist regions is most distinct in the tropics. A pair of zonal circulations rotating in the opposite direction can be seen between the Pacific Ocean and India Ocean and between Africa and Atlantic Ocean, respectively. In NH summer, double InterTropical Convergence Zones (double ITCZs) occur over the western Pacific.

Keywords Water vapor · COSMIC RO · Radiosonde · Circulation · Double ITCZs

J. Luo (✉) · Z. Chen · L. Yi
School of Geodesy and Geomatics, Wuhan University, SGG, Wuhan, China
e-mail: jluo@sgg.whu.edu.cn

Z. Chen
e-mail: 810098620@qq.com

L. Yi
e-mail: 763598311@qq.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*, Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_43,

43.1 Introduction

Water vapor is one of the most important greenhouse gases in the Earth's atmosphere. It is the key factor in the monitoring of global climate change. The distribution of water vapor is uneven over the globe. It changes rapidly with smaller scale than wind and temperature. It is still very difficult to monitor the atmospheric humidity precisely because of the rapid variations of moisture and the lack of adequate real-time observations [1]. The Global Positioning System (GPS) radio occultation (RO) technique, which in cooperation with conventional moisture detection means including radiosonde, satellite observation, water vapor radiometer, radar, aircraft detection, laser detection, spectrum analyzer and ground hygrometer detection, makes the real-time monitoring of moisture more and more easily. This helps us to understand the atmospheric water vapor variations better.

Radiosonde is regarded as one of the main traditional techniques to detect the atmospheric water vapor. It provides the most reliable observations from the Earth's Troposphere to the lower Stratosphere with long-term stability and has been the backbone for operational forecasting and a key data source for climate analysis [2, 3]. With high quality radiosondes, temperatures measured within ± 0.2 K, pressures measured within ± 0.5 hpa ($= 0.5$ mb), and relative humidity measured within ± 2 % or higher can be obtained [4, 5]. Oort [6] presented the global distribution of specific humidity with radiosonde observations firstly and made a detailed description of the vertical variability at 11 different vertical levels. In 1996, Oort and Peixoto analyzed the three dimensional distribution of relative humidity and summarized the common characteristics of water vapor distribution based on 15-year data sets [5]. The spatial distribution of those radiosonde stations is uneven. The measurements over the ocean area and over the Earth southern hemisphere (SH) are sparse, which will bring errors in the analysis of global water vapor variations. Therefore, it is hard to monitor the water vapor distribution with high capacity, near real-time and high vertical resolution with radiosondes only.

To some extent, GPS RO can make up the shortcomings of the conventional water vapor detection technique. The vertical resolution of humidity profiles derived from GPS RO observations is about 100–200 m in the lower troposphere and 1.4 Km in the stratosphere, which is higher than those of any satellite data. And GPS RO observations are of the characteristics of global coverage, long-term stability, high accuracy and all-weather capability. On 14 Apr 2006, the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC), jointly developed by United States and China's Taiwan Province was successfully launched in the Vandenberg Air Force Base with the way of "an arrow six star". COSMIC mission provides up to 2,000 RO profiles over the globe every day. Due to the advanced tracking technique, known as the 'open-looping tracking', more than 90 % of COSMIC profiles arrived the lowest 2 km of the troposphere [3]. The precision and accuracy of

GPS RO profiles has been demonstrated in many papers [7]. Kursinski and Hajj [8] compared the specific humidity derived from GPS/MET RO mission and the reanalysis data from the European Center for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) during the period from 21 June to 4 July 1995. The results show that the accuracy of specific humidity obtained from GPS RO can reach 0.2–0.5 g/Kg in the middle and lower troposphere which meets the requirements of global water vapor analysis [8]. In 2010, Ho et al. [3] compared the specific humidity profiles derived from COSMIC RO with radiosonde observations and with the reanalysis dataset from ECMWF. They found that the moisture information provided by the three data sources agree well generally. They also found that the types of radiosonde stations may have impacts on the comparison results [3]. All above studies show that GPS RO observations meet the requirements for the detection of water vapor variations.

In this study, we presents and demonstrates the accuracy of COSMIC RO specific humidity profiles through statistical comparisons between collocated RO and radiosonde data observed from 1 Jan to 31 Dec 2009. Moreover, we give the analysis of the global water vapor distribution based on COSMIC RO observations. The two datasets and the comparison methods are briefly introduced in Sect. 43.2. Section 43.3 gives the detailed results and discussion. The conclusions are given in Sect. 43.4.

43.2 Datasets and Comparison Methods

In this paper, the radiosonde data is provided by the Integrated Global Radiosonde Archive (IGRA) of the National Climate Data Center (NCDC).

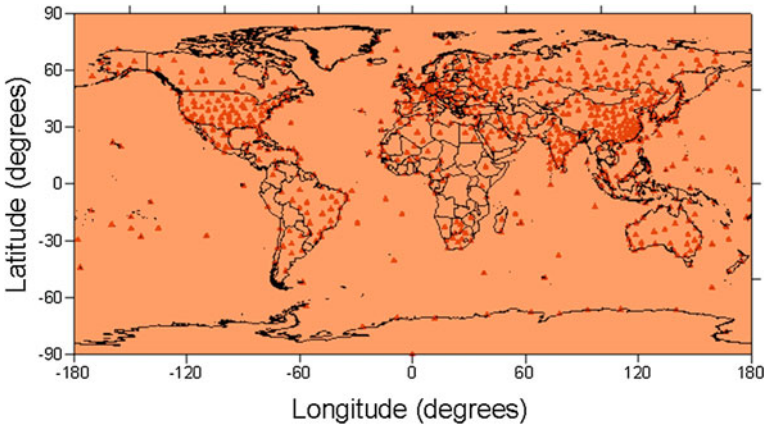


Fig. 43.1 Global distribution of radiosonde stations in 2009

The IGRA consists of radiosonde and pilot balloon observations at more than 1,500 worldwide stations with varying periods of record, many of which extend from 1960s to present. The quality of the data was checked with rigorous procedures and the reliability of the data was demonstrated by many researches [9, 10]. There are 784 operational radiosonde stations during 2009–2010 and their geographical distribution is shown in Fig. 43.1.

The COSMIC RO profiles are provided by COSMIC Data Analysis and Archive Center (CDAAC). The specific humidity (q) profiles from the GPS RO are compared with collocated radiosonde observations. The collocation criterion is as following: the maximum radial distance and the maximum time difference between COSMIC RO and radiosonde measurements is d = 300 Km and Δt = 3h, respectively [11]. In the neutral atmosphere, the relationship between specific humidity (q), pressure (p) and water vapor pressure (e) is shown by equation (43.1):

$$q = \varepsilon \cdot \frac{e}{p - (1 - \varepsilon)e} \tag{43.1}$$

where $\varepsilon = R_d/R_v (\approx 0.622)$, and R_d and R_v is the gas constant for dry air and water vapor, respectively. Equation (43.1) is used to get the specific humidity from the water vapor pressure sounded by radiosondes.

In the comparison, the COSMIC RO and the radiosonde specific humidity profiles are interpolated into the same 500 m vertical grids. For each available collocated pair:

$$\Delta q(l,m) = \frac{q_{COSMIC}(l,m) - q_{RS}(l,m)}{q_{RS}(l,m)} \tag{43.2}$$

In Eq. (43.2), $\Delta q(l,m)$ is the relative specific humidity deviation, $q_{COSMIC}(l,m)$ and $q_{RS}(l,m)$ is the COSMIC RO and the radiosonde specific humidity, respectively. l is the index for vertical height level and m is the index for the observation pair. The average of the relative specific humidity (q) deviation $\overline{\Delta q(l)}$and the standard deviation $\sigma_{\Delta q(l)}$ at each height level is computed with Eqs. (43.3) and (43.4) respectively:

$$\overline{\Delta q(l)} = \frac{1}{M(l)} \sum_{m=1}^{M(l)} \Delta q(l,m) \tag{43.3}$$

$$\sigma_{\Delta q(l)} = \left\{ \frac{1}{M(l) - 1} \sum_{m=1}^{M(l)} \left[\Delta q(l,m) - \overline{\Delta q(l)} \right]^2 \right\}^{1/2} \tag{43.4}$$

After the accuracy of the COSMIC RO specific humidity profiles is validated, the distribution of global specific humidity (q) during northern hemisphere (NH) summer (June–August) and NH winter (December–February) of 2009 are analyzed based on COSMIC RO data.

43.3 Results and Discussion

43.3.1 Comparison of Specific Humidity Profiles Between COSMIC RO and Radiosonde Data

Before analyzing the specific humidity distribution with COSMIC RO observations, we must validate the accuracy of the COSMIC RO specific humidity profiles. The radiosonde data is usually taken as the reference in this type of validation [12]. Here, the COSMIC RO profiles are collocated with the observations of PENANG/BAYAN LEPAS radiosonde station (5.30°N, 100.27°E) during each former half-day (00–12UT) of 2009. An example of such collocated pairs is shown in Fig. 43.2a, which demonstrates the comparison of specific humidity profiles between collocated COSMIC RO and radiosonde soundings. The radiosonde observation occurred at 00 UT on 10 July 2009. The collocated COSMIC RO event occurred at the location of 7.13° N, 99.20°E at 0153 UT in the same day. The collocated specific humidity profiles are shown at the height level of 2–25 km. Figure 43.2b depicts the statistic comparison of specific humidity profiles between radiosonde and COSMIC RO observations. 51 collocated pairs are got under the collocation criterion of the maximum radial distance d = 300 Km and the maximum time difference Δt = 3h. The results show that the average relative bias is less than

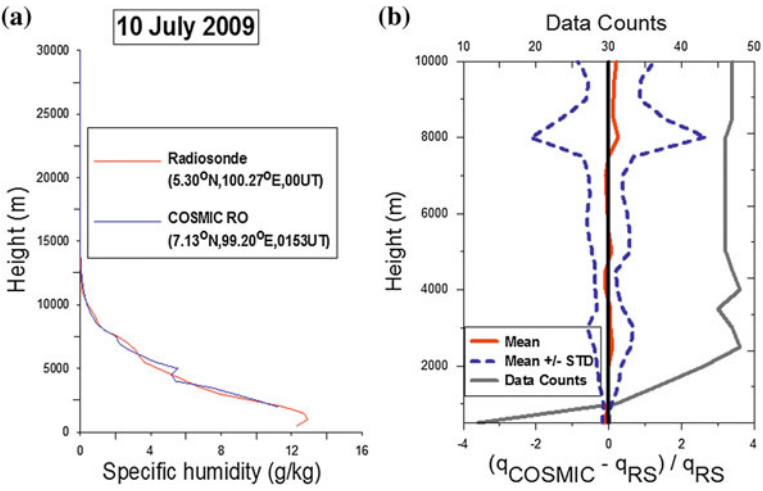


Fig. 43.2 **a** A example of collocated radiosonde and COSMIC RO specific humidity profiles. **b** Statistical comparisons of 51 pairs of collocated radiosonde and COSMIC RO specific humidity profiles under the collocation criteria of the maximum radial distance d = 300 Km and the maximum time difference Δt = 3h between COSMIC RO and radiosonde measurements

20 % at all height levels, and the standard deviation is generally less than 0.7 under the height level of 7.5 km.

43.3.2 Distribution of Global Specific Humidity

Figures 43.3 and 43.4 shows the distribution of global water vapor at different height levels during NH summer and NH winter of 2009 respectively. Since the specific humidity profiles below 1.4 Km are affected by the negative bias of refractivity in the atmospheric boundary layer, the biases of COSMIC RO observations become large and the number of RO events arriving at the surface of earth decreases in many areas, especially in the tropical regions. The vertical resolution of specific humidity is about 1.5 Km in the upper troposphere and the lower stratosphere (UTLS, 5–35 Km). We depict the specific humidity at five altitude levels with vertical resolution of 1.5 Km starting from 1.4 Km to analyze the distribution of water vapor. As shown in Figs. 43.3 and 43.4, the distribution of water vapor presents distinct gradient variation from low latitude to high latitude regions except in some special areas, such as the Sahara desert of North Africa, which reflects the upward and downward branches of the Hardley cell, Ferrel cell and the polar circulation during NH summer and NH winter. The specific humidity in the low latitudes (30°S–N) is larger than other areas. As altitude level increases, the water vapor content decreases. In addition, the contrast between dry and moist regions is most distinct over the tropics, such as the Pacific Ocean–India Ocean and Africa–Atlantic Ocean, where we can see the pair of zonal circulations rotating in the opposite direction at 1.4 and 2.8 km altitude levels, respectively. Both Figs. 43.3 and 43.4 show that the water vapor distribution over the globe is of zonal asymmetry.

From Fig. 43.3, it can be seen that the maximum value of specific humidity is got in the north–eastern part of India, which extends up to the Pacific Ocean. In the north and south of west Pacific equator, two specific humidity bands which are parallel to the equator and are of at least 15° in width are distinct. Over these two bands, the specific humidity at 1.4 Km (∼850 hpa) altitude level is larger than 10 g/kg. This meets the standards of double InterTropical Convergence Zones (double ITCZs). The specific humidity is larger over the tropics than over other areas. In contrast, the specific humidity is less than 4 g/kg in the Sahara desert of North Africa. From Fig. 43.4, it can be seen that during the NH winter, the specific humidity is large at 1.4 km (∼850 hpa) and 2.9 km (∼700 hpa) altitude levels over the South America (especially Amazon region), southern part of Africa, Malaysia, and Indonesia. Besides, there is a band with large specific humidity at all altitude levels extending from 6°N to 25°S and from 80°E to 120°E. In Figs. 43.3 and 43.4, it can also be seen that the ITCZ migrates southward in NH winter and northward in NH summer.

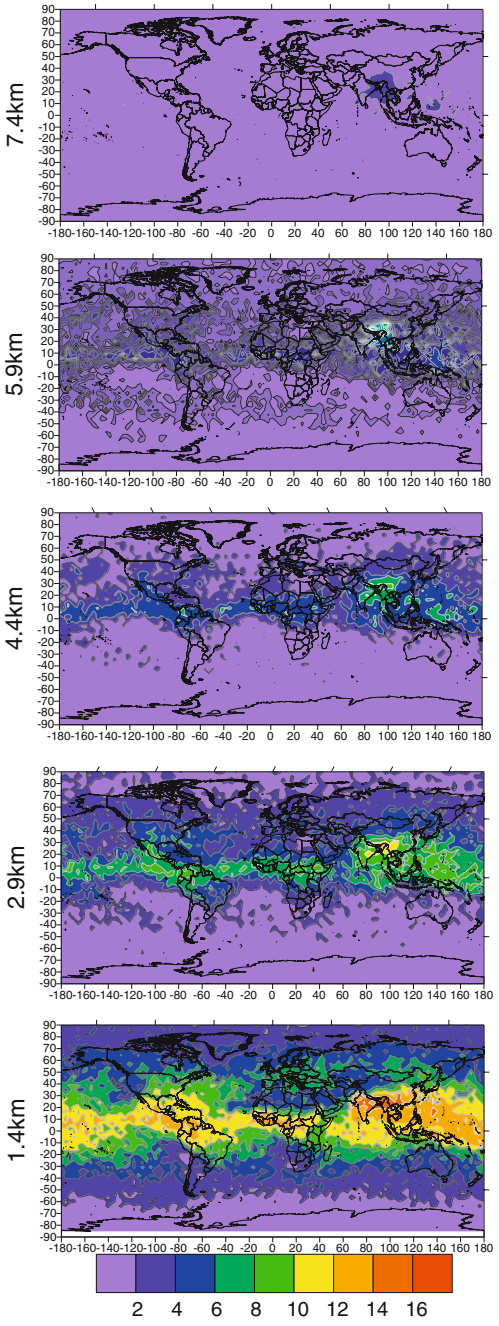


Fig. 43.3 The distribution of specific humidity during NH Summer (June–Aug) of 2009

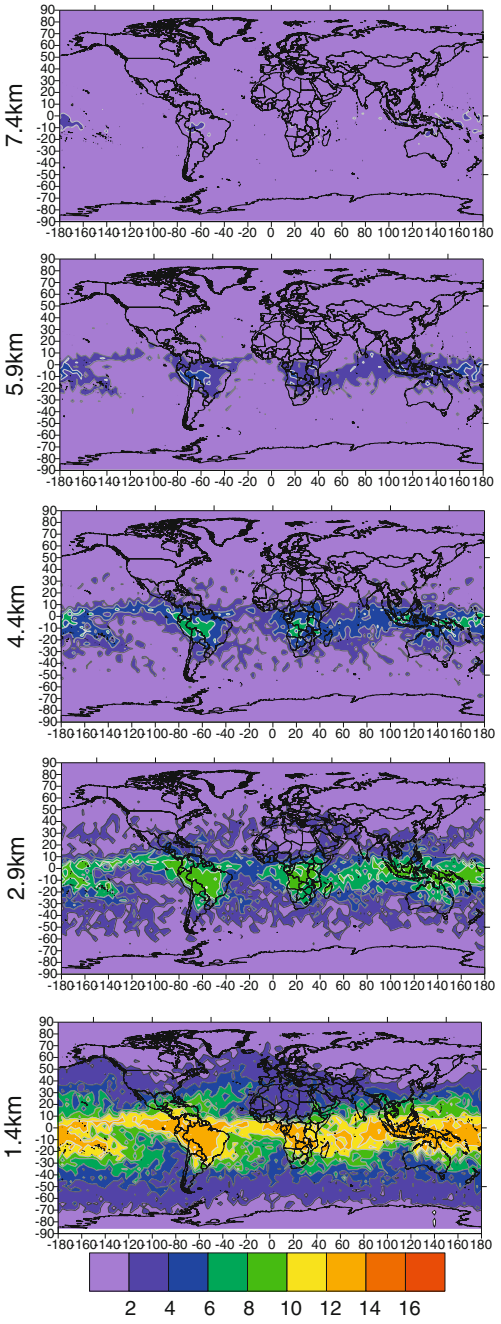


Fig. 43.4 The distribution of specific humidity during NH Winter (Dec–Feb) of 2009

43.4 Conclusions and Future Work

In this study, the COSMIC RO specific humidity profiles of 1 Jan to 31 Dec 2009 are compared with collocated radiosonde soundings provided by IGRA of NCDC. The collocation criterion is that the maximum radial distance d = 300 Km and the maximum time difference Δt = 3h. The distribution of specific humidity during NH summer and NH winter over the globe are analyzed. The following conclusions are drawn:

1. COSMIC RO specific humidity profiles agree well with radiosonde soundings of IGRA. The average relative biases are less than 20 % at all height levels. The standard deviations are less than 0.7 below 7.5 km.
2. Specific humidity decreases as the altitude increases. The distribution of specific humidity reflects some climatic phenomena such as the double ITCZs, Hardley cell, Ferrel cell and the polar circulation during NH summer and NH winter. The water vapor coverage over the globe is of zonal asymmetry. At the same time, the contrasts between dry and moist regions are most distinct over the tropics.

Acknowledgments The authors acknowledge CDAAC for providing COSMIC RO data, and NCDC for providing IGRA radiosonde data. We thank all the reviews for their critical comments and suggestions which made us to improve this paper and future research.

References

1. Li GP et al (2010) Ground-based GPS meteorology. Science Press, China. 李国平, et al., 地基GPS气象学. 科学出版社, 2010
2. Dabberdt WF et al (2002) Radiosondes. Encyclopedia of Atmospheric Sciences, Vol 6. pp 1900–1913
3. Ho S-P et al (2010) Global evaluation of radiosondewater vapor systematic biases using GPS radio occultation from COSMIC and ECMWF analysis. Remote Sens 2 (5):1320–1330
4. Wickert J (2004) Comparison of vertical refractivity and temperature profiles from CHAMP with radiosonde measurements. Danish Meteorological Institute, Denmark
5. José PP (1996) The Climatology of relative humidity in the atmosphere
6. Oort AH, S. United, Global atmospheric circulation statistics, 1958–1973 (1983) NOAA professional paper; 14. 1983, Rockville, U.S. Dept of Commerce, National Oceanic and Atmospheric Administration, Washington; For sale by the Supt of Docs, US GPO xiii, p 180
7. Steiner AK et al (2011) GPS radio occultation for climate monitoring and change detection. Radio Sci 46
8. Kursinski ER, Hajj GA (2001) A comparison of water vapor derived from GPS occultations and global weather analyses. J Geophys Res 106(D1):1113
9. Durre I, Vose RS, Wuertz DB (2005) Overview of the integrated global radiosonde archive. J Clim 19(1):53–68

10. Eskridge ER et al (1995) A comprehensive aerological reference data set (CRADS): rough and systematic errors. Bull Am Meteorol Soc 76:1759–1775
11. Xu X, Luo J, Shi C (2009) Comparison of COSMIC radio occultation refractivity profiles with radiosonde measurements. Adv Atmos Sci 26:1137–1145
12. Xu X, Zhang K, Luo J, Fu E (2009) Research on the collocation criteria in the statistical comparisons of GPS radio occultation and radiosonde observations. Geomatics Info Sci Wuhan Univ 34(11):1332–1335. 徐晓华,Zhang Kefei,罗佳,Fu Erjiang. GPS掩星与探空观测统计比较中配对标准的比较研究. Geomatics and Information Science of Wuhan University 34 (11):1332–1335
13. Kishore P et al (2011) Global (50°S–50°N) distribution of water vapor observed by COSMIC GPS RO: comparison with GPS radiosonde, NCEP, ERA-Interim, and JRA-25 reanalysis data sets. J Atmos Solar Terr Phys 73(13):1849–1860

Chapter 44
The Analysis of the Characterization for GLONASS and GPS on-Board Satellite Clocks

Wenju Fu, Guanwen Huang, Yilin Liu, Qin Zhang and Hang Yu

Abstract With GLONASS and GPS satellite clocks aging and modernizing, their characteristics present corresponding changes and certain differences. For that reason, the paper puts forward a combined method of detecting gross, adopts the quadratic polynomial model to fit the satellite clock error model, calculates the phase, frequency, frequency drift and residual of GLONASS and GPS on-board satellite clocks, chooses Modified Total Variance and Hadamard Total Variance to calculate the stability of cesium (Cs) and rubidium (Rb) clock respectively, and then analyses the characterization of the on-board satellite clocks. From examples, finding out that the combined method of detecting gross is effective; revealing the variation law of the phase, frequency, drift and residual for GLONASS and GPS satellite clocks; discovering that the stability of GPS BLOCK IIF satellite clocks are the best, GPS BLOCK IIR-M&IIR Rb clocks and the newer GLONASS satellite clocks follow, and then there are GPS BLOCK IIA Cs clocks and the older GLONASS satellite clocks, but GPS BLOCK IIA Rb clocks are the worst. Meanwhile, there is a numerical relationship between the stability and residual of the satellite clocks.

Keywords Satellite clock · Stability · Frequency · GPS · GLONASS

44.1 Introduction

In the recent years, with the modernization of GLONASS, all the GLONASS satellite clocks have been replaced by GLONASSS-M clocks. Simultaneously, GPS has stepped into the aging phase, of which GPS BLOCK IIA satellite clocks

W. Fu (✉) · G. Huang · Y. Liu · Q. Zhang · H. Yu
School of Geology Engineering and Geomatics, Chang'an University,
No.126, Yanta Road, Xi'an, China
e-mail: fuwenju2009@163.com

J. Sun et al. (eds.), *China Satellite Navigation Conference (CSNC) 2013 Proceedings*,
Lecture Notes in Electrical Engineering 244, DOI: 10.1007/978-3-642-37404-3_44,

have been running for more than 18 years. Although GPS has started to modernize, only two satellite clocks of GPS BLOCK IIF type have been launched by now [1–3]. Considering the critical role of satellite clocks in the satellite navigation system, it's necessary to evaluate the characterization of GLONASS and GPS on-board satellite clocks in order to ensure the quality and stability of the satellite navigation service.

Based on these purposes, the paper collects the 30 s-interval final precise clock error data from March 1, 2012 to December 1, 2012 provided by the European Space Operations Centre (ESOC) of European Space Agency (ESA) [4]. Firstly, the paper preprocesses the clock error, then adopts the quadratic polynomial model to fit the satellite clock error model, calculates the phase, frequency, drift and residual of GLONASS and GPS on-board satellite clocks, figures out the stability of cesium (Cs) and rubidium (Rb) clock respectively with Modified Total Variance and Hadamard Total Variance. Lastly, the paper reveals the variation law of the phase, frequency, drift and residual of GLONASS and GPS satellite clocks and achieves some significant results through examples.

44.2 Analysis Models of the Satellite Clock Characteristics

The characteristics of satellite clocks mainly include: phase, frequency, frequency drift, stability and so on. During the long-term running process of satellite clocks, the gross errors and data interruption often occur due to the influence of external environment. Therefore, it's necessary to preprocess the clock error before analysing the characteristics. The mathematical models of satellite clocks are as follows:

44.2.1 Model of Detecting Gross Error

The paper adopts the median method (MAD) to detect gross errors [5, 6]. The median MAD can be expressed as:

$$MAD = Median\{|y_i - m|/0.6745\}. \tag{44.1}$$

where m is the median of the clock error sequence, that is, $m = Median\{y_i\}$; When $|y_i - m| > n \cdot MAD$ (the integer n is determined by the need), the clock error is regarded as the gross error.

44.2.2 Satellite Clock Error Model

Because the phase, frequency, frequency drift of satellite clocks constitute three parameters of the quadratic polynomial model, Consequently, often using the quadratic polynomial to fit the satellite clock error model and then analysing these characteristics. The quadratic polynomial model is shown as follow [7, 8]:

$$\Delta t_i = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 + v_i. \tag{44.2}$$

where a_0, a_1, a_2 are the satellite clock parameters, standing for phase, frequency, frequency drift respectively; t_0 is the reference epoch of the satellite clock parameters, Δt_i is the clock error.

44.2.3 Model of Analysing the Satellite Clock Stability

The Modified Total Variance is often used to describe the stability of Cs clocks while the Hadamard Total Variance is often used to Rb clocks [9–14].

Based on the phase data, the Modified Total Variance and Hadamard Total Variance can be expressed as:

$$Mod\sigma^2_{total}(\tau) = \frac{1}{2(m\tau_0)^2(N-3m+1)} \sum_{n=1}^{N-3m+1} \left(\frac{1}{6m} \sum_{i=n-3m}^{n+3m-1} \left[\bar{x}^{\#}_i(m) - 2\left(\bar{x}^{\#}_{i+m}(m)\right) + \bar{x}^{\#}_{i+2m}(m) \right]^2 \right) \tag{44.3}$$

$$H\sigma^2_{total}(\tau) = \frac{1}{6(m\tau_0)^2(N-3m)} \sum_{n=1}^{N-3m} \left(\frac{1}{6m} \sum_{i=n-3m}^{n+3m-1} \left[x^{\#}_{i+3m} - 3x^{\#}_{i+2m} + 3x^{\#}_{i+m} - x^{\#}_i \right]^2 \right). \tag{44.4}$$

where $\tau = m\tau_0$, N is the number of phase data; $\left\{x^{\#}_i\right\}$ is the $9m + 1$ subsequence mapped by the sequence removing frequency drift. The time interval of neighbouring clock error data is $\tau \cdot \bar{x}^{\#}_n(m) = \frac{1}{m} \sum_{j=0}^{m-1} x^{\#}_{n+j}$.

44.3 Data Preprocessing

In order to compare characteristics of GLONASS and GPS satellite clocks conveniently, the following figures list their running time before December 30, 2012.

In this study, the paper explores the methods of deciding n and detecting gross error on the basis of the median method and the satellite clock error model, and further preprocesses GLONASS and GPS satellite clock error data.

From formula (44.1), *MAD* is determined by the clock error data, while n is by the need. If the value of n is small, the effective clock error will be deleted, on the contrary, the gross errors won't be deleted, which influences the evaluation of satellite clock characteristics. Consequently, the right value of n is the key of data preprocessing and the premise of assessing the satellite clock characteristics.

The paper detects gross error with the median method and fits the satellite clock error model with the quadratic polynomial model, and then determines the suitable value of the integer n based on gross error number and fitting precision, finally reaches the purpose of detecting gross error reasonably. Figure 44.1 and Table 44.1 show the gross error number and fitting precision of GLONASS and GPS satellite clocks under different limiting error using the phase data, while Fig. 44.2 shows the results of detecting gross error combining the phase data and frequency data.

From the above charts, we can see:

1. For GLONASS, starting from $n = 5$, the gross error number detected by the median method shows a slight change with the increase of n, but no longer reduces significantly, and the average fitting precision reaches 0.97 ns, which shows the effective clock error data are fully reserved and the gross errors are

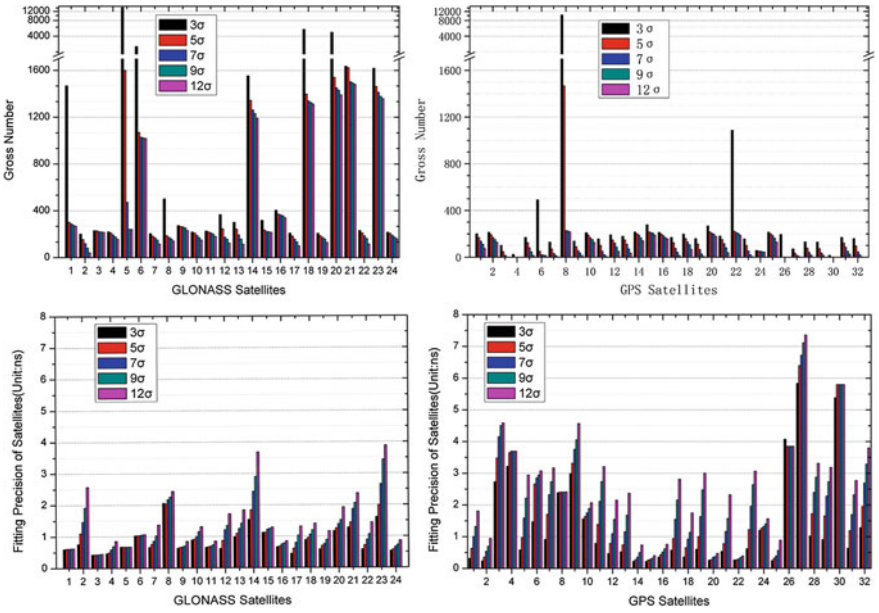


Fig. 44.1 The gross error number and fitting precision of GLONASS and GPS satellite clocks under different limiting error

Table 44.1 The average fitting accuracy of GLONASS and GPS constellation satellite under different limiting error (unit: ns)

Limiting error (σ: mean square error)	3σ	5σ	7σ	9σ	12σ
GLONASS	0.87	0.97	1.13	1.29	1.53
GPS	1.34	1.69	2.02	2.33	2.66

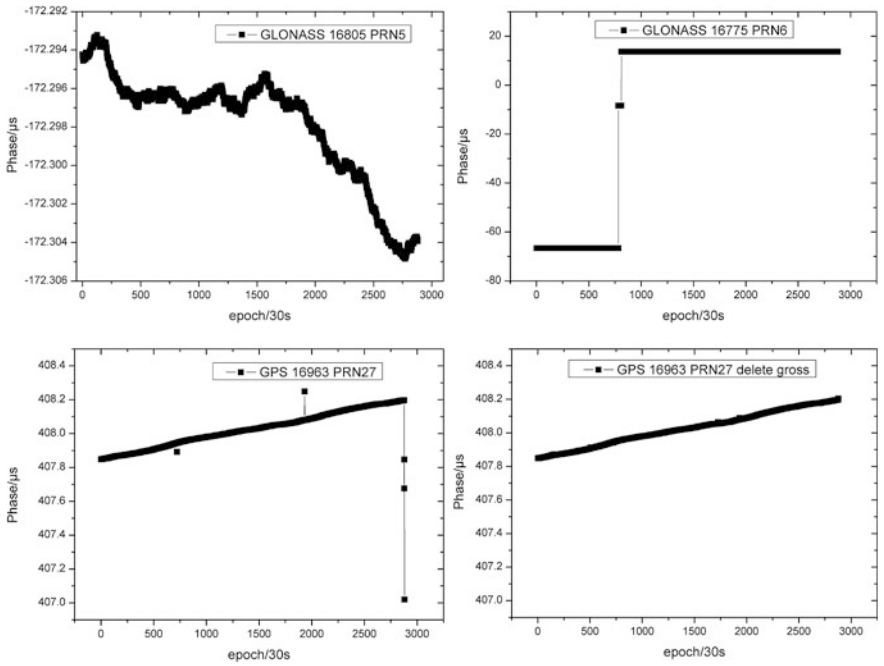


Fig. 44.2 Phase data of GLONASS 05 and 06 and GPS 27 phase data with and without gross errors

also eliminated effectively. Thus, for GLONASS, using 5 times of the mean square error as the limiting error to detect gross error; while n is 3 for GPS.

2. GLONASS 05, 06 satellite clocks of Fig. 44.2 show if the gross error number is large and shows a slight change with the increase of n, it indicates phase operation or phase skip happens; while if the gross error number is large and shows a significantly change with the increase of n, it indicates the satellite clock error data have a large fluctuation; So this can judge the satellite clock's status.
3. The advantages of using phase data to detect gross error: it can position gross error and find phase skip conveniently. Disadvantages: it can't find small gross error and frequency skip. The advantages and disadvantages of using frequency data to detect gross error are opposite. So the paper puts forward a combined method of detecting gross error by integrating their advantages, which firstly uses phase data to detect gross error, finds and processes phase skip, and then employs frequency data to detect small gross error and frequency skip. Figure 44.2 shows gross errors of GPS 27 clock error data have been fully eliminated.

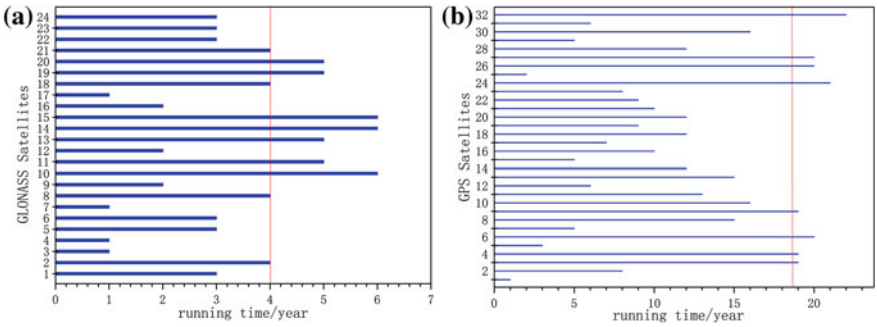


Fig. 44.3 The running time of GLONASS and GPS satellite clocks

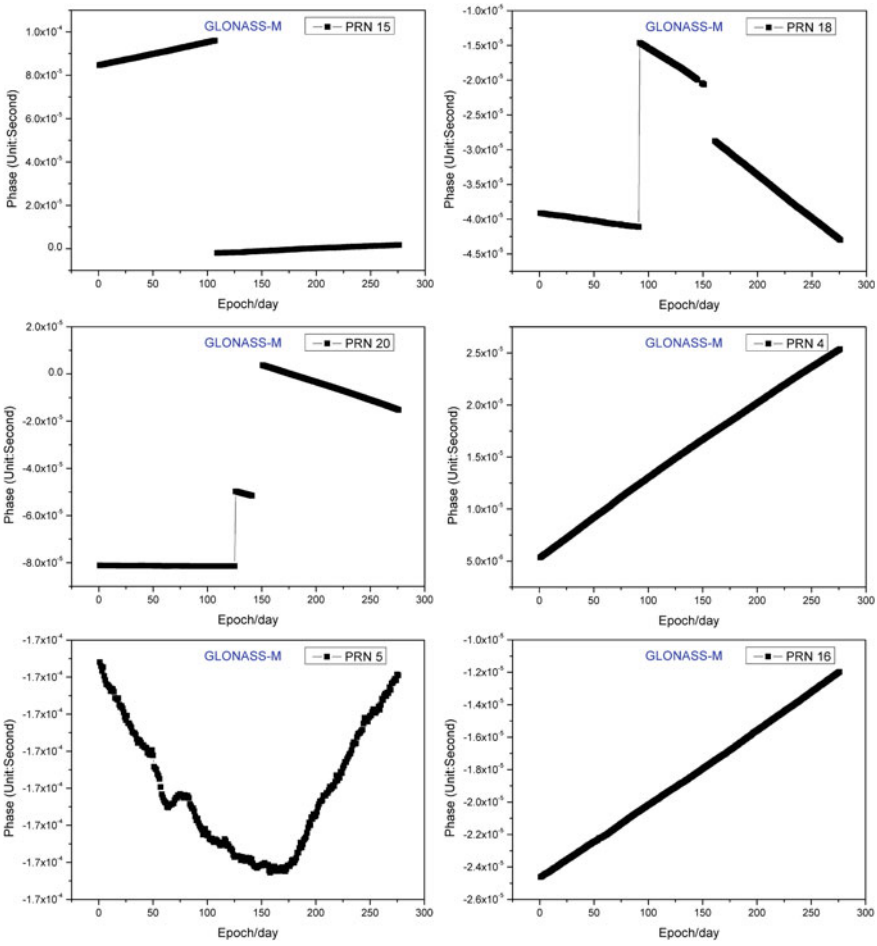


Fig. 44.4 The phase varying law of GLONASS on-board satellite clocks

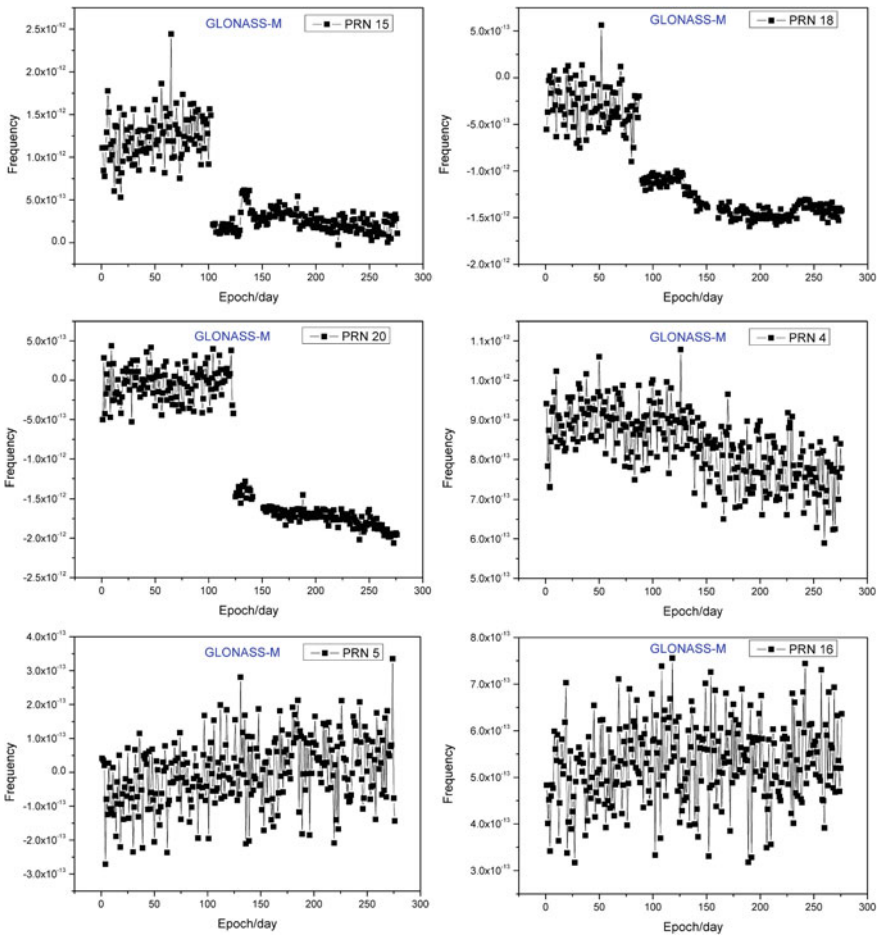


Fig. 44.5 The frequency varying law of GLONASS on-board satellite clocks

44.4 Analyses of Phase, Frequency, Drift and Residual Characteristics for Satellite Clocks

Figure 44.3 shows all the GLONASS on-board satellite clocks are GLONASS-M type whose average on-board time is 3.4 years while GPS on-board satellite clocks include early IIA and IIR type, as well as modernized IIR-M and IIF type. Among them, the average on-board time of IIA type is about 18.8 years, 10.8 and 5.3 years for IIR and IIR-M type respectively, while the satellite clocks of IIF type launch in the recent two years [2, 3]. The paper analyses the phase, frequency, frequency drift and residual of GLONASS and GPS on-board satellite clocks through adopting the quadratic polynomial model to fit the satellite clock error model for every single

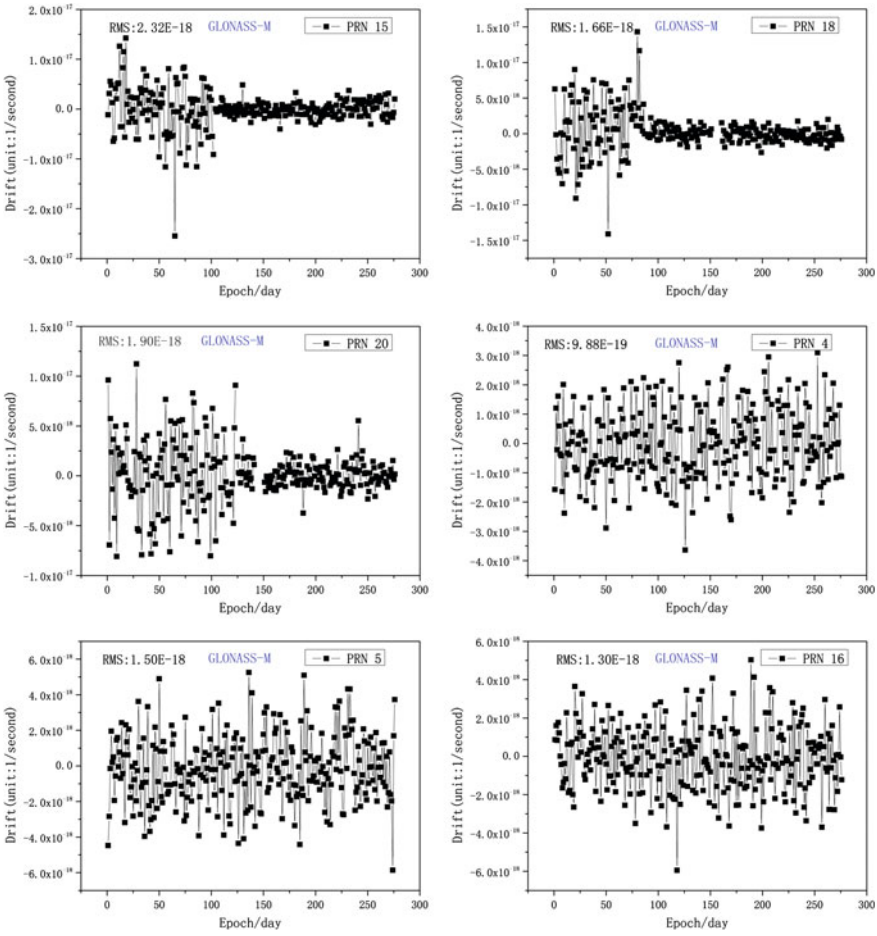


Fig. 44.6 The drift varying law of GLONASS on-board satellite clocks

day. Considering the length of paper, this paper only enumerates figures of the older GLONASS 15, 18, 20 and newer GLONASS 04, 05, 16, as well as GPS 30 of IIA, GPS 02, 16 of IIR, GPS 12, 15 of IIR-M and GPS 25 of IIF.

From the above charts, we can see:

1. From the phase, frequency, frequency drift and residual figures of GLONASS and GPS satellite clocks, the older GLONASS satellite clocks present frequent phase and frequency operations, and their frequency drift and residual diminish remarkably after these operations, such as GLONASS 15, 18, 20. While these characteristics of the newer GLONASS satellite clocks change steadily, such as GLONASS 04, 05, 16. Moreover, the early GPS IIA and IIR satellite clocks also present frequent phase and frequency operations, such as GPS 30, 02, but

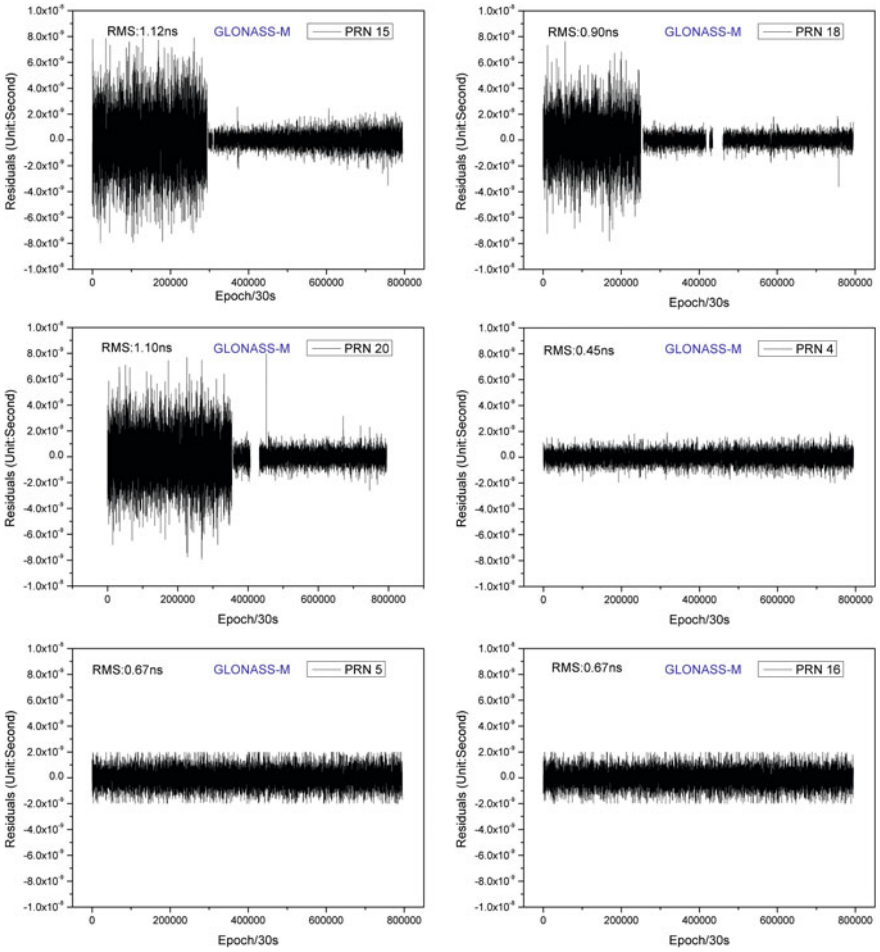


Fig. 44.7 The residual varying law of GLONASS on-board satellite clocks

the modernized GPS IIR-M and IIF type have less phase and frequency operations and their phase data are stable. And the phase and frequency operations of GPS satellite clocks have little influence on frequency drift and residual (Figs. 44.4, 44.5, 44.6, 44.7, 44.8, 44.9, 44.10, and 44.11).

2. From Figs. 44.4, 44.5, 44.6, and 44.7, GLONASS 05 is particular in its up-opening parabola-shaped phase data, linear changing frequency, frequency drift of 1.5 × 10–18, stable changing residual and fitting precision of 0.67 ns.
3. The fitting precision figures of GLONASS and GPS satellite clocks show the mean fitting precision of GLONASS and GPS satellite clocks are 0.82 and 0.93 ns. The fitting precision of the older GLONASS satellite clocks are obviously worse than the newer GLONASS satellite clocks. The aging GPS IIA

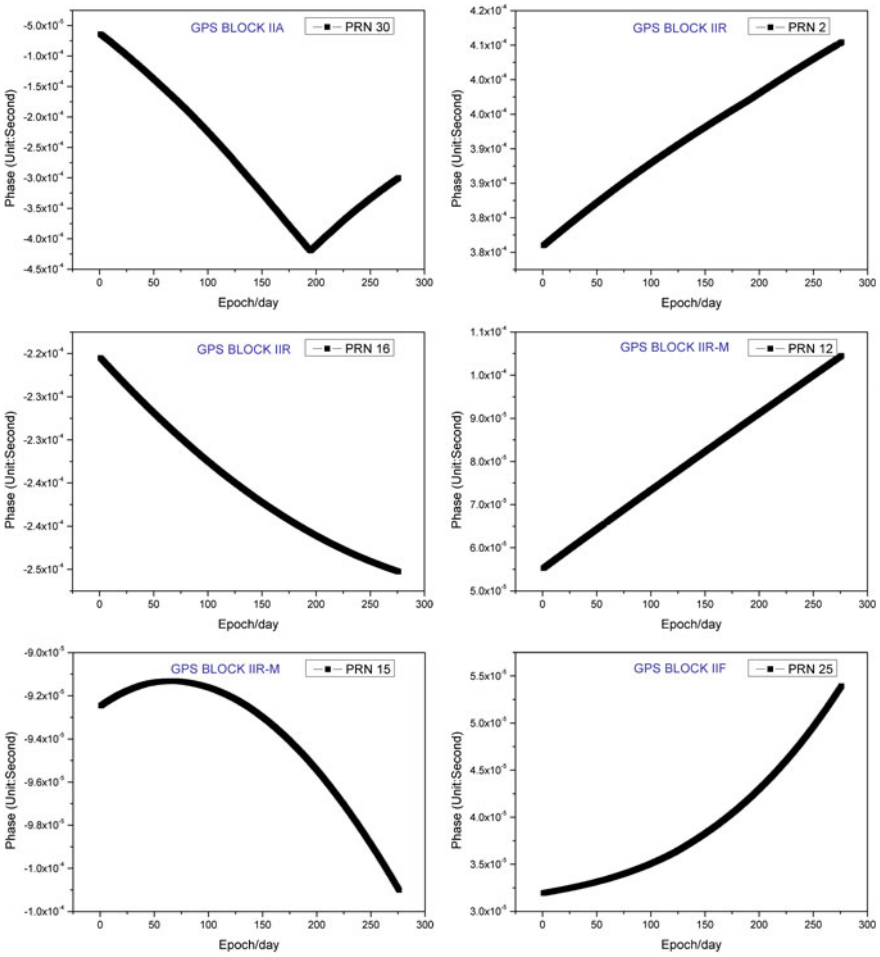


Fig. 44.8 The phase varying law of GPS on-board satellite clocks

satellite clocks are obviously worse than modernized GPS satellite clocks, and worse than GLONASS satellite clocks. However, the modernized GPS satellite clocks are obviously superior to the GLONASS satellite clocks (Fig. 44.12).

44.5 Analyses of the Stability Characteristics for Satellite Clocks

Now, GLONASS on-board satellite clocks are Cs clock, while GPS clocks are mainly Rb clock and only five satellite clocks of early IIA are Cs clock. Because Rb clocks have obvious frequency drift, and the Modified Total Variance and

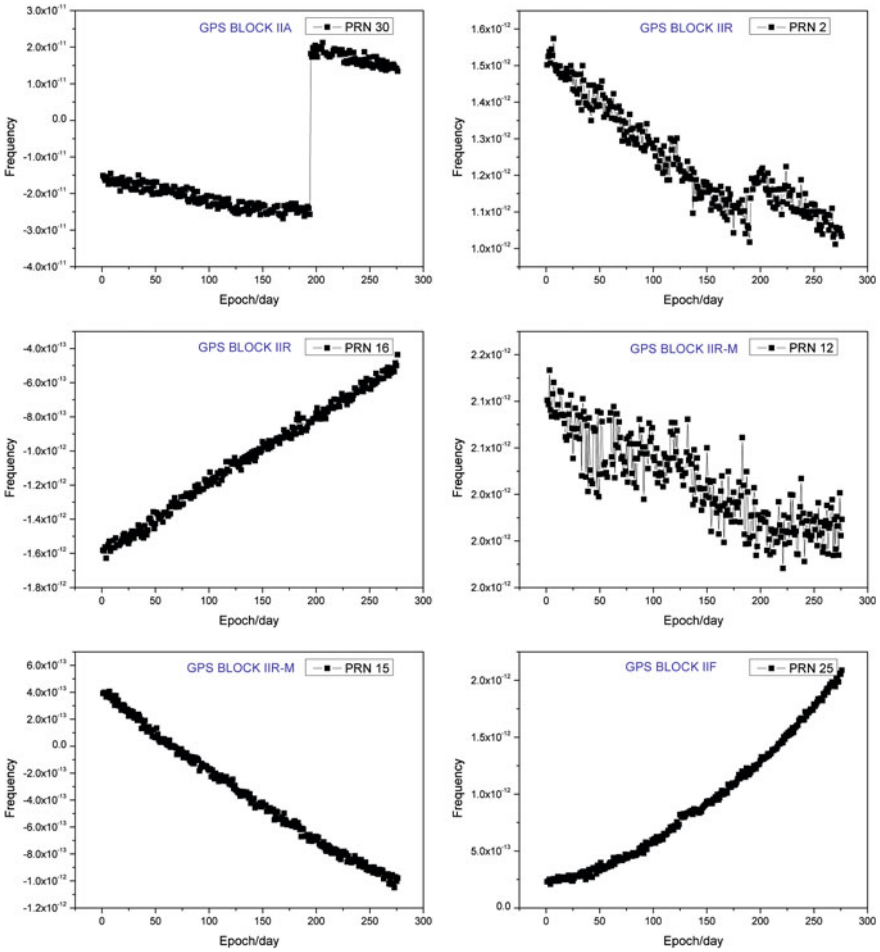


Fig. 44.9 The frequency varying law of GPS on-board satellite clocks

Hadamard Total Variance have higher confidence, therefore, the paper adopts Modified Total Variance and Hadamard Total Variance to describe the stability of Cs and Rb clock respectively. Considering we often use the 30 s and 5 min precise satellite clock error product, therefore, this paper analyses their short-term stability variation law of GLONASS and GPS satellite clocks. Because of the paper's length, only the 30 s short-term stability variation law figures of GLONASS and GPS satellite clocks are shown as follow (Figs. 44.13, 44.14, and 44.15).

In order to reveal the stability variation law of GLONASS and GPS satellite clocks further, in view of the lack of several satellite clock error data, the paper chooses the 30 s-interval final precise clock error data from November 16, 2012 to November 20, 2012 to analysis the stability of GLONASS and GPS satellite clocks under different smoothing time. For increasing the readability, the paper divides

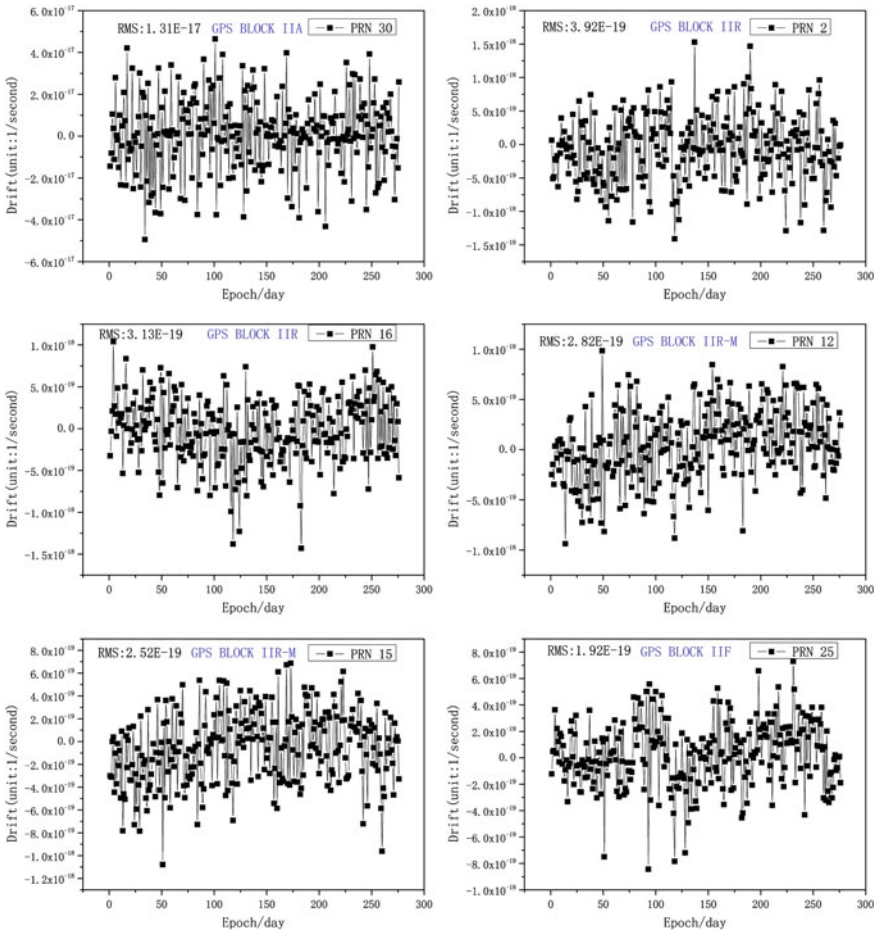


Fig. 44.10 The drift varying law of GPS on-board satellite clocks

GLONASS and GPS satellite clocks to four groups: GLONASS clock; GPS IIA Cs clock; GPS IIA and IIR Rb clock; GPS IIR-M and IIF Rb clock.

From the above charts, we can see:

1. The variation law of residual and stability shows there is a numerical relation between the residual and stability. For examples, when the stability of GLONASS 15, 18, 20 decrease, the residual decrease accordingly. This confirms the law described by the Ref. [1].
2. The longer the running time of GLONASS and GPS satellite clocks is, the worse their stability gets. For example, the stability of GLONASS 06 and 08 launched in 2009 and 2008 respectively are worse, while the stability of GLONASS 17 launched in November 2011 is better. Another example is that

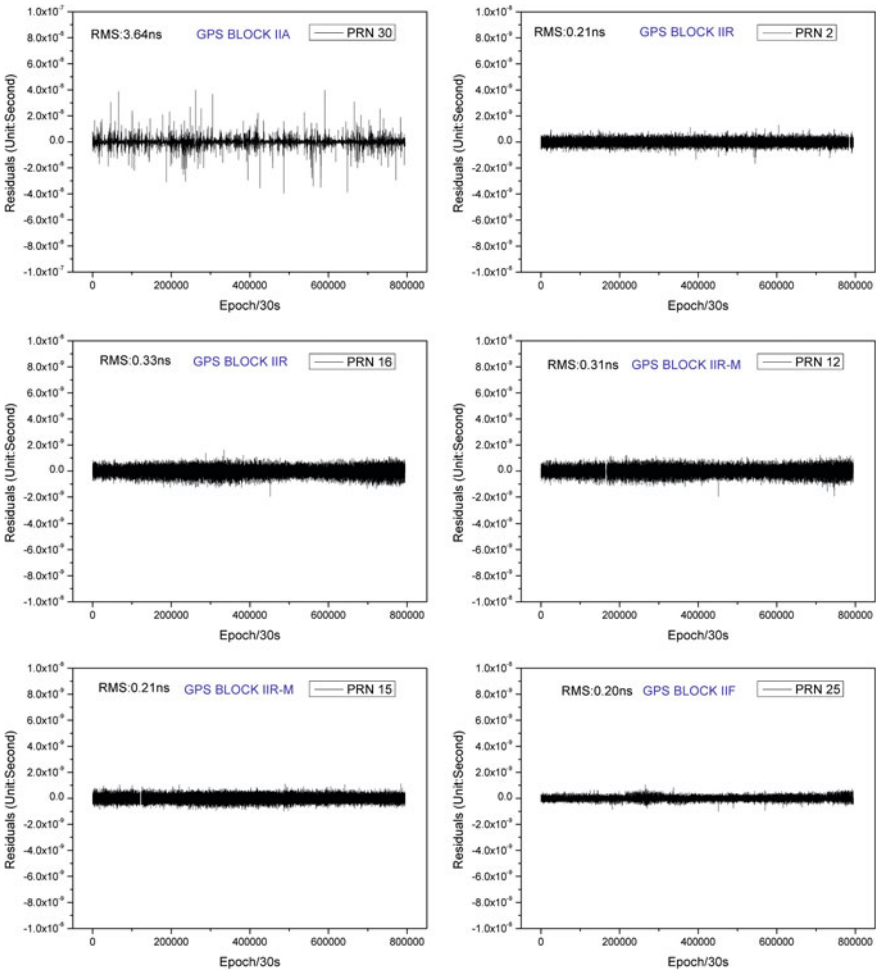


Fig. 44.11 The residuals varying law of GPS on-board satellite clocks

the average running time of GPS IIA satellite clocks is about 18.8 years, so their Rb clock's stability are the worst obviously.

3. The stability of GPS IIF satellite clocks are the best, GPS IIR-M&IIR Rb clocks and the newer GLONASS satellite clocks follow, and then there are GPS IIA Cs clocks and the older GLONASS satellite clocks, but GPS IIA Rb satellite clocks are the worst. From Figs. 44.16 and 44.17, the stability of GPS 01, 25 are the best, GPS 12, 14, 19 and GLONASS 17 follow, and then there are GPS 08, 10 and GLONASS 06, 08, 23, but GPS 04, 30 are the worst. However, the stability of some of the newer GLONASS satellite clocks and GPS IIF satellite clocks are almost in the same level, so are most of the older GLONASS satellite clocks and GPS IIR-M and IIR satellite clocks. Such as GPS 01, 25 and

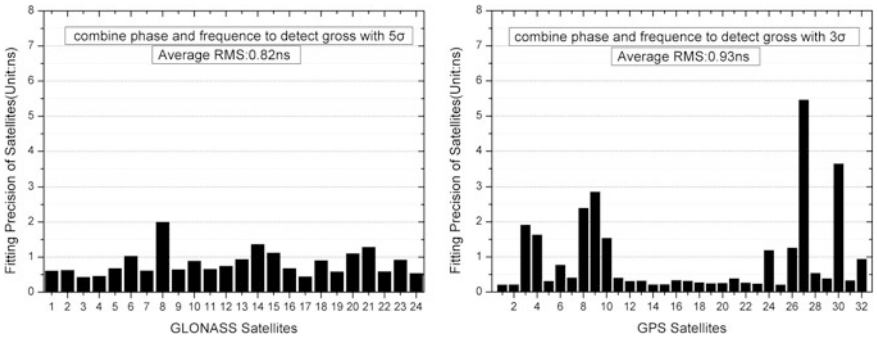


Fig. 44.12 The fitting precision of GLONASS and GPS satellite clocks without gross errors

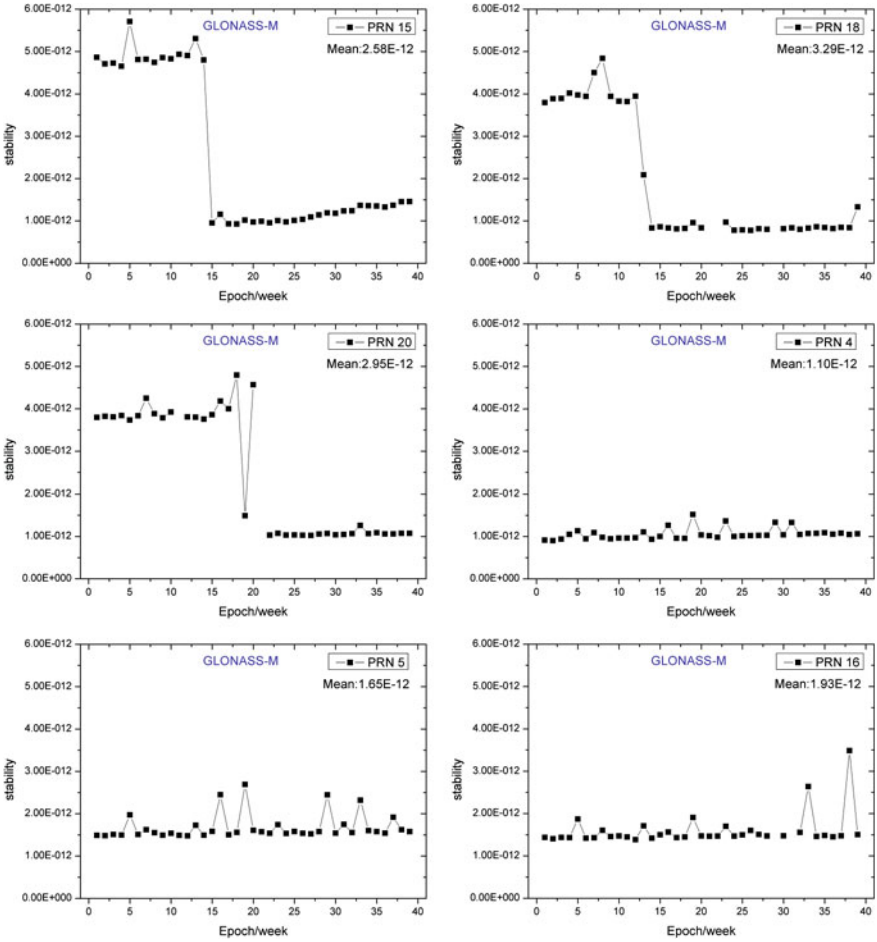


Fig. 44.13 The short-term stability variation law of the GLONASS satellite clocks

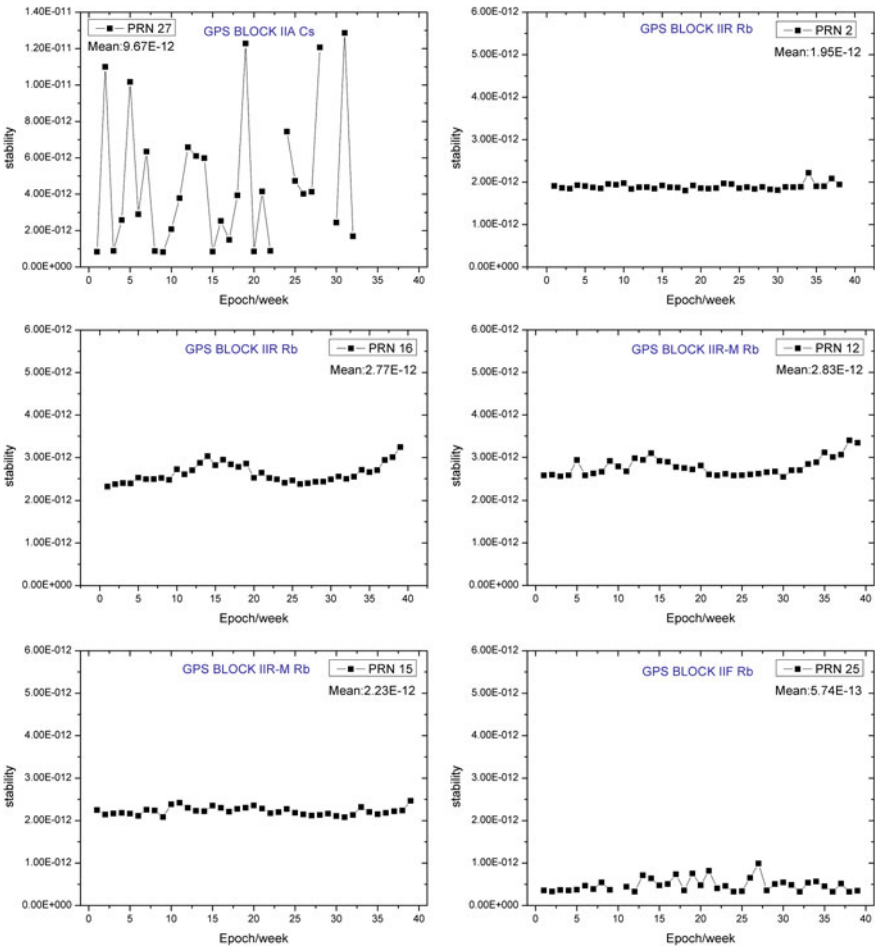


Fig. 44.14 The short-term stability variation law of the GPS satellite clocks

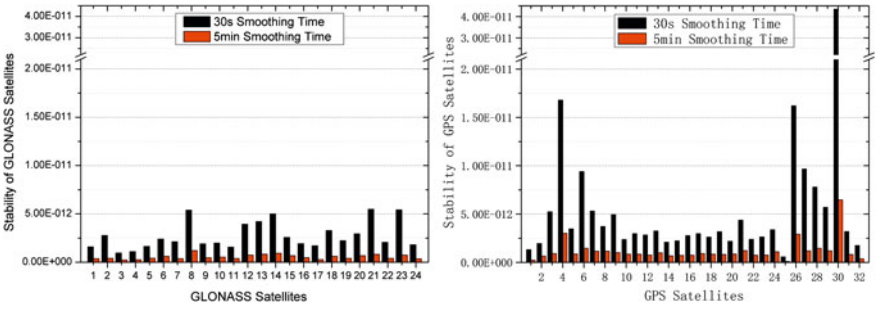


Fig. 44.15 The short-term stability of the GLONASS and GPS satellite clocks

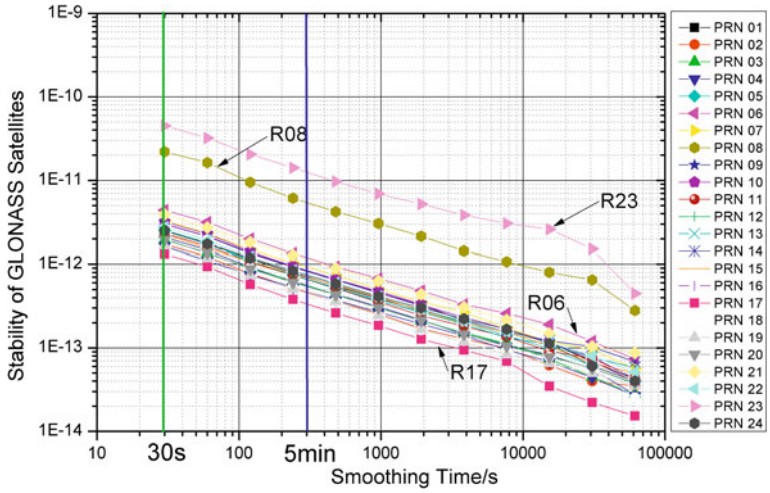


Fig. 44.16 The stability of the GLONASS satellite clocks under different smoothing time

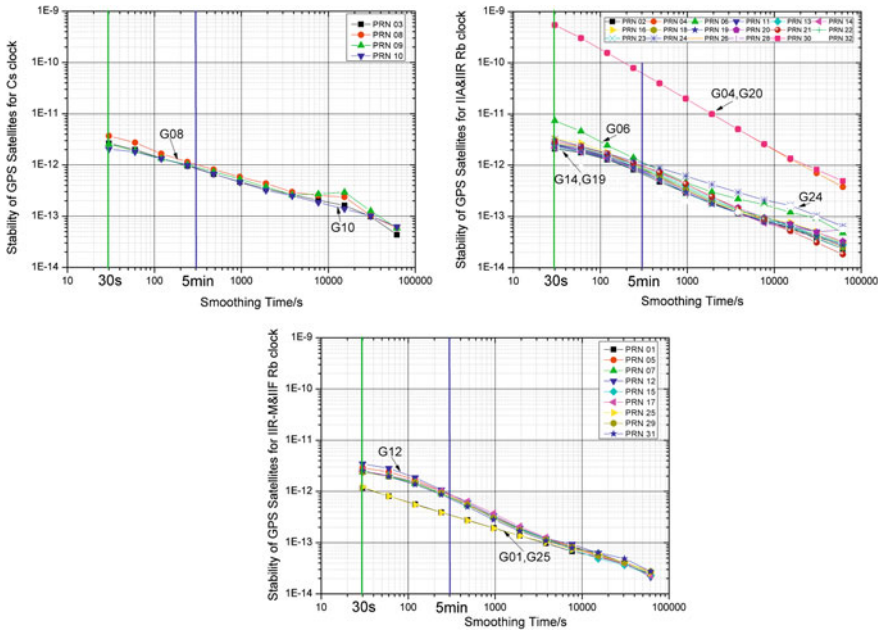


Fig. 44.17 The stability of the GPS satellite clocks under different smoothing time

GLONASS 17 are almost in the same level. Moreover, the 30 s and 5 min short-term stability of GLONASS and GPS satellite clocks are marked with green and blue vertical lines respectively in Figs. 44.16 and 44.17.

44.6 Conclusions and Suggestions

Through analysing and comparing the phase, frequency, frequency drift, residual and stability variation law of GLONASS and GPS on-board satellite clocks, the paper achieves some conclusions and suggestions:

1. Putting forward a combined method of gross detection, which can both delete gross errors effectively and find phase and frequency skip conveniently. For GLONASS, using 5 times of the mean square error as the limiting error to detect gross error is suitable; while n is 3 for the GPS.
2. The older GLONASS and GPS IIA&IIR satellite clocks present frequent phase and frequency operations, while the newer GLONASS and GPS IIR-M&IIF satellite clocks show seldom. Moreover, the frequency drift and residual of GLONASS satellite clocks diminish remarkably after these operations, while GPS satellite clocks behave normal.
3. The fitting precision of the older GLONASS satellite clocks are obviously worse than the newer. The aging GPS IIA satellite clocks are obviously worse than the modernized, and worse than the GLONASS satellite clocks. However, the fitting precision of the modernized GPS satellite clocks are superior to GLONASS satellite clocks.
4. The stability of GPS IIF satellite clocks are the best, GPS IIR-M&IIR Rb clocks and the newer GLONASS clocks follow, and then there are GPS IIA Cs clocks and the older GLONASS clocks, but GPS IIA Rb clocks are the worst. In addition, the longer the running time of GLONASS and GPS satellite clocks is, the worse their stability gets. And there is a numerical relation between the residual and stability.

Acknowledgments Thanks to the European Space Operation Centre (ESOC) of the European Space Agency (ESA) for providing the free downloading GLONASS and GPS clock error data. This study is supported by the National Natural Science Fund Project (40902081, 41104019, 40841021), the Chinese National Programs for High Technology Research and Development (2007AA12Z331), the Fundamental Research Funds for the Central Universities (CHD2010ZY001) and the Special Foundation for Chang'an university Basic Research Support Plan.

References

1. Huang G, Zhang Q, Li H, Fu W (2012) Research on quality variation of GPS satellite clocks on-orbit using IGS clock products, Adv Space Res 51:978–987. doi:10.1016/j.asr.2012.09.041
2. http://gge.unb.ca/Resources/GLONASSConstellationStatus.txt
3. http://gge.unb.ca/Resources/GPSConstellationStatus.txt
4. ftp://dgn6.esoc.esa.int/products

5. Guo H (2006) Study on the analysis theories and algorithms of the time and frequency characterization for atomic clocks of navigation satellites. Information Engineering University, ZhengZhou
6. Yang Y, Song L, Xu T (2002) Robust estimator for correlated observations based on bifactor equivalent weights. J Geodesy 76(6–7):353–358
7. Huang G, Zhang Q (2012) Real-time estimation of satellite clock offset using adaptively Robust Kalman filter with classified adaptive factors. GPS Solution 16:531–539. doi:10.1007/s10291-012-0254-z
8. Huang G, Zhang Q et al (2012) Real-time clock offset prediction with an improved GPS satellite clocks model. GPS Solution, Pressing
9. Allan DW (1966) Statistics of atomic frequency standards. Proc IEEE 54(2):221–230
10. Baugh RA (1971) Frequency modulation analysis with the hadamard variance. In: Proceedings of annual symposium on frequency control. Ultrasonics Ferroelectrics and Frequency Control Society, Roma, pp 222–225
11. Mccaskill TB, Reid WG, Largay MM, Buisson JA (1984) On-orbit frequency stability analysis of the GPS NAVSTAR-I quartz clock and the NAVSTARs-6 and -8 rubidium clocks. In: Proceedings of the 16th annual precise time and time interval (PTTI) applications and planning meeting. U. S. Naval Observatory, Washington, pp 27–29
12. Hutsell ST (1996) Operational use of the Hadamard variance in GPS. In: Proceedings of 28th PTTI meeting. U. S. Naval Observatory, Reston, pp 201–213
13. Howe D (2000) A total estimation of Hadamard function used for GPS operation. In: Proceedings of 32nd PTTI MEETING. U. S. Naval Observatory, Reston, pp 255–268
14. Hesselbarth A, Wanninger L (2008) Short-term stability of GNSS satellite clocks and its effects on precise point positioning. In: Proceedings of ION GNSS 2008, Savannah, GA, pp 1855–1863